Analysis of Channel-Based User Authentication by Key-Less and Key-Based Approaches

Stefano Tomasin

Department of Information Engineering, University of Padova, Italy

arXiv:1705.03430v2 [cs.IT] 14 Jul 2017

Abstract

User authentication (UA) supports the receiver in deciding whether a message comes from the claimed transmitter or from an impersonating attacker. In cryptographic approaches messages are signed with either an asymmetric or symmetric key, and a source of randomness is required to generate the key. In physical layer authentication (PLA) instead the receiver checks if received messages presumably coming from the same source undergo the same channel. We compare these solutions by considering the physical-layer channel features as randomness source for generating the key, thus allowing an immediate comparison with PLA (that already uses these features). For the symmetric-key approach we use secret key agreement, while for asymmetric-key the channel is used as entropy source at the transmitter. We focus on the asymptotic case of an infinite number of independent and identically distributed channel realizations, showing the correctness of all schemes and analyzing the secure authentication rate, that dictates the rate at which the probability that UA security is broken goes to zero as the number of used channel resources (to generate the key or for PLA) goes to infinity. Both passive and active attacks are considered and by numerical results we compare the various systems.

Index Terms

Physical Layer Authentication; Physical Layer Security; Rayleigh Fading; User Authentication.

I. Introduction

User authentication (UA) methods in communication systems are used to confirm the identity of a message sender [1]. In particular, the receiver must take a decision on who has transmitted the message considering that an attacker aims at impersonating the legitimate transmitter. Typically, UA includes two phases: an *identification association (ID-A)* phase, when the legitimate transmitter is assigned an identifying feature using an authenticated channel, and an *identification verification (ID-V)* phase, when the identifying feature is verified upon reception of a message.

Many commonly-used UA protocols for communication systems use as identifier a key, i.e., a secret know by either only the transmitter (asymmetric key) or both the transmitter and the receiver (symmetric key), and encryption techniques are adopted in the ID-V phase [1]. These methods go also under the name of key-based UAs [2], [3]. An alternative key-less approach is physical layer authentication (PLA) [4], in which the identifying feature is the physical channel over which the communication occurs. In this case the ID-A phase consists in the identification of the channel features, while the ID-V phase consists in checking if the received message has undergone the same channel of the ID-A phase (see e.g. [5]–[7] and [4] for a survey). For example, in wireless systems, the propagation phenomena (e.g., fading) are associated to the specific position of both transmitter and receiver, thus the attacker should be in the same position of the legitimate transmitter for a successful impersonation. PLA has been studied for discrete memory-less channels in [8]; implementations for multiple-input multiple-output (MIMO) and intersymbol-interference channels have been proposed in [5] and [9], while a game-theoretic study of PLA has appeared in [10]. Recently, a review of key-based and key-less security approaches with a comparison of cryptography and physical-layer security solutions has been proposed in [11]. In [8] the keyless UA was studied in a new noisy model, where there are two discrete memoryless channels, one from sender to receiver and and the other the from adversary to the receiver, in addition to an insecure noiseless channel between the legitimate parties. Optimality of the scheme was further proofed in [12].

In both key-based and key-less approaches a source of randomness is needed to either generate the keys or identify the user feature. In this paper we focus on the use of physical-layer channel features as randomness source for all the UA techniques. While the exploitation of channel features is intrinsic in PLA, for key-based solutions we extract a random number (the key) from the channel. The key must remain secret to the attacker, in order to prevent impersonation: thus for symmetric-key systems we use the secret key agreement (SKA) procedure of [13], [14], while for asymmetric-key systems the channel is used as an entropy source for one user only, and a set of secret bits is obtained from this source.

Most of the literature has considered separately key-based and key-less UA. In [15] a first comparison of the approaches when the unique source of randomness was the channel has been proposed, but limited to quantized channel model with a finite number of samples per frame. In this paper instead we focus on the asymptotic case of an infinite number of available independent and identically distributed (i.i.d.) channel realizations. In this asymptotic regime all considered

UA schemes can be shown to be correct, i.e., authentic messages are always accepted. We then analyze the secure authentication rate (SAR), that dictates the rate at which the probability that UA security is broken goes to zero as the number of channel resources used to extract either the key or the channel features goes to infinity. In particular, we focus on time-invariant channels, time-variant channels and Rayleigh fading channels with additive white Gaussian noise (AWGN). Most derivations are obtained considering a passive attacker that listens to transmissions in the ID-A phase and then attempts to impersonate the legitimate transmitter in the ID-V phase. However, we also consider active attacks aiming at disrupting the ID-A phase in order to ease the impersonation attack in the ID-V phase. Beyond considering optimal attacks by Eve we also discuss a simple approach in which the attacker first performs a linear combination of all channel estimates to obtain the minium mean square error (MMSE) linear estimate of the legitimate channel which is then used for the attack. By numerical results we closely compare the various systems, highlighting their potentials and vulnerabilities.

The rest of the paper is organized as follows. Section II introduces the system model, with emphasis on both the channel and the considered schemes. The analysis of the protocols in terms of SAR is performed in Section III, while its derivation for the relevant case of reciprocal Rayleigh fading AWGN channels is discussed in Section IV. Active attacks during the ID-A phase are discussed in Section V. Numerical results comparing the various strategies are presented in Section VI, before the main conclusions are outlined in Section VII.

Notation: $\mathbb{P}[\mathcal{X}]$ denotes the probability of the event $\mathcal{X}$. $\mathbb{E}[x]$ denotes the expectation of the random variable x, $\mathbb{H}(x)$ denotes the entropy of the discrete random variable x, and $\bar{\mathbb{H}}(x)$ denotes the differential entropy of the continuous random variable x. $\mathbb{I}(x;y)$ denotes the mutual information between random variables x and y. $\mathbb{D}(p||q)$ denotes the Kullback-Leibler (KL) divergence between the two probability density function (PDF) p and q. Vectors are denoted in boldface, while scalar quantities are denoted in italic. $\ln x$ and $\log x$ denote the natural-base and base-2 logarithms of x, respectively. $\det \boldsymbol{A}$ and $\mathrm{tr}\boldsymbol{A}$ denote the determinant and trace of matrix $\boldsymbol{A}$, respectively. $p_x(a)$ denotes the PDF of random variable x.

II. System Model

We consider a communication system with two legitimate users, Alice and Bob, and one attacker Eve. Time is divided into frames, each of the same duration. In order to simplify the analysis we suppose that in even and odd frames Alice and Bob alternate their transmissions. In

particular, starting from the first frame and in all odd frames $2t+1$ ($t=0,\ldots$) Alice transmits to Bob, while starting from the second frame and in all even frames $2t$ ($t=1,\ldots$) Bob transmits to Alice. Eve can transmit at any frame.

Bob must decide if packets received in odd frames are coming from either Alice or Eve. To this end, initial frames are used for ID-A purposes, extracting keys from the channel or identifying reference channel features as described in the following. In forthcoming frames Bob receives packets and performs ID-V, i.e., he decides if they are coming from Alice (hypothesis $\mathcal{H}_0$) or not (hypothesis $\mathcal{H}_1$).

Channel between each users' couple are assumed to be time-invariant within each frame, and described by n complex random variables. All transmissions include pilot symbols known to all users (including Eve) for channel estimation. The channel estimated by Bob at frame $2t+1$ is described by the random vector

$$\boldsymbol{X}(2t+1)=[x_1(2t+1),\ldots,x_n(2t+1)], \tag{1}$$

while the channel estimated by Alice in frame $2t$ is described by the random vector

$$\boldsymbol{Y}(2t)=[y_1(2t),\ldots,y_n(2t)]. \tag{2}$$

Let

$$\boldsymbol{V}_{\mathrm{A}}(2t+1)=[v_{1,\mathrm{A}}(2t+1),\ldots,v_{n,\mathrm{A}}(2t+1)] \tag{3}$$

$$\boldsymbol{V}_{\mathrm{B}}(2t)=[v_{1,\mathrm{B}}(2t),\ldots,v_{n,\mathrm{B}}(2t)] \tag{4}$$

be the estimated channels by Eve when either Alice or Bob are transmitting, respectively. We also collect into the vector $\boldsymbol{Z}(2t)$ the channel estimates of Eve up to frame $2t$, i.e.,

$$\boldsymbol{Z}(2t)=\left[\boldsymbol{V}_{\mathrm{A}}(1),\boldsymbol{V}_{\mathrm{B}}(2),\ldots,\boldsymbol{V}_{\mathrm{A}}(2t-1),\boldsymbol{V}_{\mathrm{B}}(2t)\right]. \tag{5}$$

Eve will exploit all these estimates to perform her attacks.

Channel estimates $\boldsymbol{X}(t_1)$, $\boldsymbol{Y}(t_2)$, $\boldsymbol{V}_{\mathrm{A}}(t_3)$ and $\boldsymbol{V}_{\mathrm{B}}(t_4)$ are assumed to be correlated, in general however they do not have the same value since they are affected by noise, time-varying channel phenomena and changing positions of transmitters and receivers. Within each estimated vector, we assume that $x_k(t)$, are i.i.d. for $k\in[1,n]$, with frame-dependent PDF $p_{x(t)}(a)$, $a\in\mathbb{C}$. Similarly we have that $y_k(t)$, $v_{k,\mathrm{A}}(t)$ and $v_{k,\mathrm{B}}(t)$ are i.i.d.. In the following we will drop index

k from all channel estimates, i.e., using $x(t)$ instead of $x_k(t)$ (for any k). Moreover, we define the $2t$-size vector

$$\boldsymbol{z}(2t) = [v_{\mathrm{A}}(1), v_{\mathrm{B}}(2), \ldots, v_{\mathrm{A}}(2t-1), v_{\mathrm{B}}(2t)]\,. \tag{6}$$

We consider three UA approaches, namely asymmetric channel-based cryptographic authentication (A-CBCA), symmetric channel-based cryptographic authentication (S-CBCA) and physical layer authentication (PLA). While both A-CBCA and S-CBCA have a cryptography background and exploit some secret known by either or both the legitimate parties, PLA has information-theoretic foundations and mostly relies on the communication channel characteristics. In order to obtain a fair comparison of the three approaches, we extract the secret needed for both A-CBCA and S-CBCA from the channel, so that all three schemes have the same source of randomness.

Moreover, both Alice and Bob must be able to exchange authenticated messages in the ID-A phase, to ensure that they are talking to each other, and not with Eve. This is a common feature to all UA protocols and we assume it for granted. The authenticated channel can be obtained for example when devices are in a protected location immune from attacks. Also, the authenticated channel is available when the keys used for authentication must be renewed, but we can still use the old keys in the ID-A phase. Moreover PLA, which is not based on a key, can be considered as a low-complexity alternative to be used in replacement of a more expensive UA procedure that is adopted only for the ID-A phase of the PLA protocol.

We now details the UA procedures for the three authentication approaches.

A. A-CBCA

The basic structure of the A-CBCA can be summarized as follows.

ID-A Phase: In this phase Alice generates a private and public key couple from a source of randomness, where the public key decrypts messages encrypted with the private key. Then, she transmits the public key to Bob over an authenticated channel.

ID-V Phase: In this phase, whenever Alice wants to transmit a message, she encrypts it with her private key, including a signature. Bob decrypts the message with the public key and checks if the signature is correct, and in this case the message is accepted as authentic, otherwise it is discarded as non-authentic.

Suitable variants of this scheme have been proposed to prevent various attacks (including the replay attack) and are out the scope of this paper. With proper adaptations, the scheme is proofed

to be secure [1], provided that the private key is secret.

In our implementation of A-CBCA, the random bits are obtained from the channel. In particular, in the second frame Alice extracts random bits from $\boldsymbol{Y}(2)$. On her side, Eve attempts to extract the private key from her channel observations. Therefore, we must ensure that the randomness extraction procedure provides bit that are not known to Eve. Proof of correctness and security in this case is provided in Section III.

B. S-CBCA

The basic structure of the S-CBCA can be summarized as follows.

ID-A Phase: Alice and Bob share over an authenticated and confidential (to Eve) channel a secret key.

ID-V Phase: Alice encrypts the message with her secret key, including a signature. Bob decrypts the message with the same secret key and checks if the signature is correct.

Also this scheme, with proper modifications and improvements (out of the scope of this work) is proofed to be secure [1], provided that the key is secret. In our implementation of S-CBCA the key is obtained from the channel, using what is known as a SKA protocol [16] that provides a secret key to two parties, with the support of an authenticated channel. In particular, in the first two frames they estimate their channels and using a public error-free authenticated channel they complete the reconciliation process. The proof of correctness and security in this case is provided in Section III.

C. PLA

With PLA the authentication is provided by the physical channel over which the communication occurs [5]. The PLA algorithm works as follows.

ID-A Phase: In this phase, occurring at frame 1, Bob obtains a reference estimate of the channel to Alice $\boldsymbol{X}(1)$. The estimation must be performed using an authenticated message, so that Bob is sure that the estimated channel is connecting him to Alice (rather than Eve).

ID-V Phase: In this phase, occurring for frames $2t+1$, with $t > 0$, whenever Bob receives a message, he decides that it comes from Alice if the estimated channel in that frame, $\boldsymbol{X}(2t+1)$, is similar to $\boldsymbol{X}(1)$.

The correctness and security proof of PLA is provided in Section III.

III. Protocols Analysis With ID-V Attacks

In this section we consider only the attacks in ID-V phase, while attacks in the ID-A phase will be considered in Section V. We assess the correctness and security (as usually defined in cryptography) of each UA protocol. We first provide a definition of correctness and security:

Definition 1 (Correctness and Security): A UA protocol is correct when a message coming from Alice is verified as authentic by Bob and is secure when a message coming from Eve is dismissed as non-authentic by Bob.

Let $\boldsymbol{K}(n)$ be the key used for key-based UA or the features used for PLA: since we assume that the n channel realizations are i.i.d., the length of $\boldsymbol{K}(n)$ grows linearly with n. We aim at establishing the protocols' correctness and security asymptotically, as the number of channel uses per frame n goes to infinity. About security, we show that for all schemes, the probability of successful attack (PSA) $P_s(t)$ at frame t goes exponentially to zero as n goes to infinity and we define

$$R(t) = \lim_{n\to\infty} -\frac{1}{n}\log P_s(t) \tag{7}$$

as the *secure authentication rate (SAR)* at frame t. Interpreting the n entries of the channel estimate vectors as n *channel uses*, $R(t)$ is the number of *secret bits in* $\boldsymbol{K}(t)$ *per channel use* at frame t, i.e, the number of bits that are not known to Eve in $\boldsymbol{K}(n)$, as $n \to \infty$. Note that the PSA depends on the frame index t since Eve collects channel estimates as time goes on, thus being able to refine her knowledge of the Alice-Bob channel. In the following we derive the maximum SAR for each UA protocol achieved by proper use of channel knowledge. We consider attacks at odd frames, starting from frame $t = 3$, after Eve has collected an even number of channel observations.

A. S-CBCA

We establish the correctness and security of the S-CBCA protocol by the following Theorem.

Theorem 1: As $n \to \infty$, the S-CBCA protocol is correct and the SAR for frame $2t+1$, with $t \geq 1$ is

$$R_{\mathrm{S-CBCA}}(2t+1) = C_{\mathrm{SKA}}(2t)\,, \tag{8}$$

where $C_{SKA}(2t)$ is the *weak secret-key capacity* of the SKA process between Alice and Bob, given the channel knowledge by Eve at frame $2t$.

Proof: S-CBCA is proofed to be correct and secure [1] provided that the key is secret. About correctness, asymptotically SKA ensures that Alice and Bob have the same key, as long as the secret key rate $R_{\mathrm{SKA}}(2t+1)$ satisfies $R_{\mathrm{SKA}}(2t+1) \leq C_{\mathrm{SKA}}(2t+1)$.

About security, for SKA the length of the secret key (in bits) grows as $nR_{\mathrm{SKA}}(2t+1)$. Therefore

$$P_s(t) \approx 2^{-nR_{\mathrm{SKA}}(2t+1)} \tag{9}$$

and from (7) the maximum SAR coincides with $C_{\mathrm{SKA}}(2t+1)$, and we obtain (8). ■

We recall that the secret key capacity for the source-model SKA is not known with a closed-form expression, but is bounded as

$$\begin{aligned} \mathbb{I}(x(1);y(2)) - \min\{\mathbb{I}(x(1);\boldsymbol{z}(2t)), \mathbb{I}(y(2);\boldsymbol{z}(2t))\} \\ \leq C_{SKA}(2t+1) \leq \\ \min\{\mathbb{I}(x(1);y(2)), \mathbb{I}(x(1);y(2)|\boldsymbol{z}(2t))\}, \end{aligned} \tag{10}$$

since the two channel estimates used for SKA are $\boldsymbol{X}(1)$ and $\boldsymbol{Y}(2)$ and the information on the channel available at Eve at frame $2t+1$ is $\boldsymbol{Z}(2t)$.

B. A-CBCA

For A-CBCA we first consider the case in which the channel estimates are discrete-valued random variable. This occurs for example if Alice quantizes $y(2)$ into $\langle y(2) \rangle \in \mathcal{A}$, where $\mathcal{A} = \{a_0, \ldots, a_{M-1}\}$ is an M-size quantization alphabet. The random number used for the private key is then extracted from $\langle y(2) \rangle$.

Theorem 2: The A-CBCA scheme is correct and for discrete-valued estimated channels $\langle y(2) \rangle$, as $n \to \infty$ its SAR at frame $2t+1$, with $t \geq 1$ is

$$R_{\mathrm{A-CBCA}}(2t+1) = \mathbb{H}(\langle y(2) \rangle | \boldsymbol{z}(2t)), \tag{11}$$

where $\mathbb{H}(\cdot|\cdot)$ is the conditional entropy.

Proof: Correctness of A-CBCA scheme follows immediately by the assumption that the public-key broadcast is error-free and authenticated.

About security, the A-CBCA protocol can be seen as a source-based SKA protocol where both Alice and Bob observe the same source of randomness. In this case the SKA capacity upper

bound becomes

$$\begin{aligned}&\min\{\mathbb{I}(\langle y(2)\rangle;\langle y(2)\rangle),\mathbb{I}(\langle y(2)\rangle;\langle y(2)\rangle|\boldsymbol{z}(2t))\}\\&=\mathbb{I}(\langle y(2)\rangle;\langle y(2)\rangle|\boldsymbol{z}(2t))=\mathbb{H}(\langle y(2)\rangle|\boldsymbol{z}(2t))\,,\end{aligned}\tag{12}$$

which coincides with the lower bound

$$\begin{aligned}&\mathbb{I}(\langle y(2)\rangle;\langle y(2)\rangle)-\mathbb{I}(\langle y(2)\rangle;\boldsymbol{z}(2t))\\&=\mathbb{H}(\langle y(2)\rangle)-\mathbb{H}(\langle y(2)\rangle)+\mathbb{H}(\langle y(2)\rangle|\boldsymbol{z}(2t))\,,\end{aligned}\tag{13}$$

providing (11). ■

For the channel quantization case a lower bound on SAR is obtained as follows

$$\begin{aligned}R_{\mathrm{A-CBCA}}(2t+1)&=\mathbb{H}(\langle y(2)\rangle)-\mathbb{I}(\langle y(2)\rangle;\boldsymbol{z}(2t))\\&\geq\max\{0,\mathbb{H}(\langle y(2)\rangle)-\mathbb{I}(y(2);\boldsymbol{z}(2t))\}\,,\end{aligned}\tag{14}$$

where the first line is obtained by the definition of mutual information and the second line by the data processing inequality, where we upper-bounded the mutual information on the quantized variable by the corresponding mutual information on the continuous-valued channel estimate.

Now, we consider the case in which the channel is a continuous random random variable, for which the SAR is established by the following Lemma.

Lemma 1: For continuous channel estimates at Alice (i.e., $y(2)$ is a continuous random variable), the SAR is

$$R_{\mathrm{A-CBCA}}(2t+1)=\infty\,.\tag{15}$$

Proof: We first observe that the conditional entropy in (11) can be written as

$$\mathbb{H}(\langle y(2)\rangle|\boldsymbol{z}(2t))=\int\mathbb{H}(\langle y(2)\rangle|\boldsymbol{z}(2t)=\boldsymbol{b})p_{\boldsymbol{z}(2t)}(\boldsymbol{b})\mathrm{d}\boldsymbol{b}\,.\tag{16}$$

Now, the continuous random variable $y(2)$ is the asymptotic case of the quantized $\langle y(2)\rangle$ with a number of quantization points $M\to\infty$. Asymptotically, the conditional entropy for each value of $\boldsymbol{z}(t)$ is the *limiting density of discrete points* which tends to infinity logarithmically with M under very mild assumptions [17]. Therefore, the average of the logarithm, which corresponds to the conditional entropy tends to infinity, i.e., $\mathbb{H}(\langle y(2)\rangle|\boldsymbol{z}(2t))\to\infty$, providing (15). ■

C. PLA

We establish the correctness and security of PLA by the following Theorem.

Theorem 3: The PLA protocol is asymptotically correct, and as $n \to \infty$ its SAR at frame $2t+1$, with $t \geq 1$, is

$$R_{\mathrm{PLA}}(2t+1) = \mathbb{D}(p_{x(1),x(2t+1)|\mathcal{H}_0} || p_{x(1),x(2t+1)|\mathcal{H}_1}), \tag{17}$$

where $p_{x(1),x(2t+1)|\mathcal{H}_0}$ and $p_{x(1),x(2t+1)|\mathcal{H}_1}$ are the joint PDF of $x(1)$ and $x(2t+1)$ when Alice and Eve are transmitting, respectively in the ID-V phase (frame $2t+1$).

Proof: About correctness, using the Chernoff-Stein Lemma [18, Theorem 11.8.3] we have that for $n \to \infty$ the probability of correctly authenticating messages coming from Alice can be made arbitrarily small.

About security, still from the Chernoff-Stein Lemma [18, Theorem 11.8.3] the PSA goes to zero as

$$P_s \sim 2^{-n\mathbb{D}(p_{x(1),x(2t+1)|\mathcal{H}_0} || p_{x(1),x(2t+1)|\mathcal{H}_1})}, \tag{18}$$

which directly provides (17). ■

Note that the SAR depends on the attack that Eve is performing through $p_{x(1),x(2t+1)|\mathcal{H}_1}$. Therefore it is interesting to have results under the hypothesis that Eve performs a specific attack, as in the following Lemma.

Lemma 2: If Eve is able to induce any channel estimate $\boldsymbol{X}(2t+1)$ to Bob when attacking, and assuming that Eve generates the induced estimate randomly distributed according to the PDF of the legitimate channel given Eve's observations, $p_{x(2t+1)|\boldsymbol{z}(2t),\mathcal{H}_0}$, then the SAR is upper-bounded by

$$R_{\mathrm{PLA}}(2t+1) \leq \mathbb{I}(x(1); x_{\mathcal{H}_0}(2t+1)|\boldsymbol{z}(2t)). \tag{19}$$

Proof: See Appendix A. ■

From this lemma we observe that if the statistics of estimated channels at both Alice and Bob are the same ($p_x = p_y$) then $R_{\mathrm{PLA}}(2t+1) \leq \mathbb{I}(x(1); y(2)|\boldsymbol{z}(2t))$, i.e., the PLA has the same upper-bound of S-CBCA.

Remark 1: For time-varying channels we assumed that Eve can induce any channel estimate when performing the attack. This however does not hold as the channel are changing and Eve does not know the exact channel (with respect to Bob) on which she attacks, therefore in practice she will not be able to induce any desired channel estimate. Therefore our assumption is conservative and the obtained SAR is a lower bound on the effective SAR. Moreover, for

PLA the time-varying channel scenario is particularly challenging, since the variations decrease the ability of the receiver to identify the channel in the ID-V phase. In other words, for PLA SAR will decrease over time. In particular, as $t \to \infty$ the KL divergence will tend to zero, thus nulling the SAR.

Remark 2: We have assessed the performance for the three UA methods, and from the theorems we can conclude that A-CBCA has an unlimited SAR as we increase the number of quantization bits. For PLA we have an explicit expression for SAR, depending however on the attack by Eve. For the SAR of S-CBCA we have only bounds. Moreover, for a particular attack strategy by Eve the SAR of PLA share the same upper bound of S-CBCA.

IV. Rayleigh AWGN Reciprocal Channels

We now consider a scenario in which the estimates are corrupted by AWGN, which corresponds for example to a massive MIMO system in which all users (Alice, Bob and Eve) have $\sqrt{n}$ antennas each, so that the resulting channel matrices have $n \to \infty$ entries[1]. Moreover, channel matrix entries are assumed i.i.d., and zero-mean unitary power complex Gaussian (ZMUPCG), in accordance with the Rayleigh fading model. By reordering the n entries of each matrix into a vector we obtain the channel model described in Section II, where $x(t)$, $y(t)$ and $z(t)$ are Gaussian distributed.

We assume that the channel between any couple of devices is reciprocal, therefore we have

$$x(t) = h(t) + \sigma_{\mathrm{x}} w_{\mathrm{x}}(t)\,, \quad y(t) = h(t) + \sigma_{\mathrm{y}} w_{\mathrm{y}}(t)\,, \tag{20}$$

where $w_{\mathrm{x}}(t), w_{\mathrm{y}}(t)$ are jointly ZMUPCG and $\sigma^2_{\mathrm{x}}, \sigma^2_{\mathrm{y}}$ are the noise powers at the receivers. We also define the correlation of $h(t)$ over time as

$$\mathbb{E}[h(t)h^*(t+\ell)] = \rho(\ell)\,. \tag{21}$$

Eve's channels are correlated with coefficients $\alpha_{\mathrm{A}} \in [-1, 1]$ and $\alpha_{\mathrm{B}} \in [-1, 1]$ to $h(t)$, and affected by AWGN with powers $\sigma^2_{\mathrm{v,A}}$ and $\sigma^2_{\mathrm{v,B}}$ (typically $\sigma^2_{\mathrm{v,A}} = \sigma^2_{\mathrm{v,B}}$). Therefore, her estimate of the channel to Alice is

$$v_{\mathrm{A}}(t) = \alpha_{\mathrm{A}} h(t) + \sqrt{1 - \alpha^2_{\mathrm{A}}}\, q_{\mathrm{A}}(t) + \sigma_{\mathrm{v,A}} w_{\mathrm{v,A}}(t)\,, \tag{22}$$

[1]Other antennas configurations can be considered, leading to similar results and expressions as those derived in this section. Also, orthogonal frequency division multiplexing (OFDM) systems can be cast in this model.

while her estimate of the channel to Bob is

$$v_{\rm B}(t) = \alpha_{\rm B} h(t) + \sqrt{1-\alpha_{\rm B}^2} q_{\rm B}(t) + \sigma_{\rm v,B} w_{\rm v,B}(t)\,, \tag{23}$$

where $w_{\rm v,A}(t)$, $w_{\rm v,B}(t)$ are ZMUPCG independent with respect to $h(t)$, $q_{\rm A}(t)$ and $q_{\rm B}(t)$, and $q_{\rm A}(t)$, $q_{\rm B}(t)$ are ZMUPCG variables. Note that by these definitions all channel estimates have unitary variance in the absence of noise.

A-CBCA: As we have already seen, using directly the continuous-valued channel estimate we obtain that the SAR of A-CBCA is infinite, thus we focus here on the channel quantization case. The conditional entropy in (11) can be written (by definition) as

$$\begin{aligned} R_{\rm A-CBCA}(2t+1) = \mathbb{H}(\langle y(2)\rangle | \boldsymbol{z}(2t)) = \mathbb{H}(\langle y(2)\rangle | \boldsymbol{z}(2t)) = \\ - \sum_{i=0}^{M-1} \int p_{\langle y(2)\rangle|\boldsymbol{z}(2t)}(a_i|\boldsymbol{b}) p_{\boldsymbol{z}(2t)}(b) \log p_{\langle y(2)\rangle|\boldsymbol{z}(2t)}(a_i|\boldsymbol{b}) {\rm d}\boldsymbol{b}\,. \end{aligned} \tag{24}$$

About the bound (14), recall that for a Gaussian vector $\boldsymbol{v}$ of size k with correlation matrix $\boldsymbol{R}_{\rm v}$ the differential entropy is $\bar{\mathbb{H}}(\boldsymbol{v}) = \log\det((\pi e)^k \boldsymbol{R}_{\rm v})$. Let us define

$$\begin{aligned} &\boldsymbol{R}_{[y(2),\boldsymbol{z}(2t)]} = \\ &= \mathbb{E}[[y(2), \boldsymbol{z}(2t)]^H [y(2), \boldsymbol{z}(2t)]] \\ &= \begin{bmatrix} 1+\sigma_{\rm y}^2 & \boldsymbol{r}_{\rm y}^H \\ \boldsymbol{r}_{\rm y} & \boldsymbol{R}_{\boldsymbol{z}(2t)} \end{bmatrix}, \end{aligned} \tag{25}$$

where $\boldsymbol{r}_{\rm y} = \mathbb{E}[y(2)\boldsymbol{z}^H(2t)]$ has entries

$$[\boldsymbol{r}_{\rm y}]_\ell = \begin{cases} \alpha_{\rm A}^* \rho(\ell-2) & \ell \text{ even}\,, \\ \alpha_{\rm B}^* \rho(\ell-2) & \ell \text{ odd}\,. \end{cases} \tag{26}$$

Then we have

$$\begin{aligned} &\mathbb{I}(y(2); \boldsymbol{z}(2t)) = \bar{\mathbb{H}}(y(2)) + \bar{\mathbb{H}}(\boldsymbol{z}(2t)) \\ &- \bar{\mathbb{H}}([y(2), \boldsymbol{z}(2t)]) = \log \frac{(1+\sigma_{\rm y}^2)\det \boldsymbol{R}_{\boldsymbol{z}(2t)}}{\det \boldsymbol{R}_{[y(2),\boldsymbol{z}(2t)]}} \\ &= -\log\left(1 - \frac{\boldsymbol{r}_{\rm y}^H \boldsymbol{R}_{\boldsymbol{z}(2t)}^{-1} \boldsymbol{r}_{\rm y}}{1+\sigma_{\rm y}^2}\right), \end{aligned} \tag{27}$$

where $\boldsymbol{R}_{\boldsymbol{z}(2t)} = \mathbb{E}[\boldsymbol{z}^H(2t)\boldsymbol{z}(2t)]$, with entries

$$[\boldsymbol{R}_{\boldsymbol{z}(2t)}]_{m,n} = \begin{cases} \alpha_{\mathrm{A}}^*\alpha_{\mathrm{B}}\rho(m-n) & n \text{ odd, } m \text{ even} \\ \alpha_{\mathrm{A}}\alpha_{\mathrm{B}}^*\rho(m-n) & n \text{ even, } m \text{ odd} \\ |\alpha_{\mathrm{B}}|^2\rho(m-n) & n \text{ and } m \text{ even, } m \neq n \\ |\alpha_{\mathrm{A}}|^2\rho(m-n) & n \text{ and } m \text{ odd, } m \neq n \\ 1+\sigma_{\mathrm{z}}^2 & n = m\,. \end{cases} \tag{28}$$

S-CBCA: For S-CBCA we have the bound (10). In particular for the Gaussian case we have

$$\begin{aligned} \mathbb{I}(x(1);y(2)) &= \mathbb{H}(x(1)) + \mathbb{H}(y(2)) - \mathbb{H}(x(1),y(2)) \\ &= -\log\left(1 - \frac{|\mathbb{E}[x(1)y^*(2)]|^2}{(1+\sigma_{\mathrm{x}}^2)(1+\sigma_{\mathrm{y}}^2)}\right) \\ &= -\log\left(1 - \frac{\rho(1)}{(1+\sigma_{\mathrm{x}}^2)(1+\sigma_{\mathrm{y}}^2)}\right)\,, \end{aligned} \tag{29}$$

and analogously to (27)

$$\begin{aligned} \mathbb{I}(x(1);\boldsymbol{z}(2t)) =& \log\frac{(1+\sigma_{\mathrm{x}}^2)\det\boldsymbol{R}_{\boldsymbol{z}(2t)}}{\det\boldsymbol{R}_{[x(1),\boldsymbol{z}(2t)]}} \\ &= -\log\left(1 - \frac{\boldsymbol{r}_{\mathrm{x}}^H\boldsymbol{R}_{\boldsymbol{z}(2t)}^{-1}\boldsymbol{r}_{\mathrm{x}}}{1+\sigma_{\mathrm{x}}^2}\right) \end{aligned} \tag{30}$$

where

$$\begin{aligned} &\boldsymbol{R}_{[x(1),\boldsymbol{z}(2t)]} \\ =&\mathbb{E}[[x(1),\boldsymbol{z}(2t)]^H[y(2),\boldsymbol{z}(2t)]] \\ =&\begin{bmatrix} 1+\sigma_{\mathrm{x}}^2 & \boldsymbol{r}_{\mathrm{x}}^H \\ \boldsymbol{r}_{\mathrm{x}} & \boldsymbol{R}_{\boldsymbol{z}(2t)} \end{bmatrix}\,, \end{aligned} \tag{31}$$

$$[\boldsymbol{r}_{\mathrm{x}}]_\ell = [\mathbb{E}[x(1)\boldsymbol{z}^H(2t)]]_\ell = \begin{cases} \alpha_{\mathrm{A}}^*\rho(\ell-1) & \ell \text{ even}\,, \\ \alpha_{\mathrm{B}}^*\rho(\ell-1) & \ell \text{ odd}\,. \end{cases} \tag{32}$$

Moreover we have

$$\mathbb{I}(x(1);y(2)|\boldsymbol{z}(2t)) = \mathbb{I}(x(1);y(2),\boldsymbol{z}(2t)) - \mathbb{I}(x(1);\boldsymbol{z}(2t)) \tag{33}$$

$$\mathbb{I}(x(1); y(2), \boldsymbol{z}(2t)) = \log \frac{(1+\sigma_{\mathrm{x}}^2) \det \boldsymbol{R}_{[y(2),\boldsymbol{z}(2t)]}}{\det \boldsymbol{R}_{[x(1),y(2),\boldsymbol{z}(2t)]}} \tag{34}$$

and

$$\boldsymbol{R}_{[x(1),y(2),\boldsymbol{z}(2t)]} = \begin{bmatrix} 1+\sigma_{\mathrm{x}}^2 & \rho(1) & \boldsymbol{r}_{\mathrm{x}}^H \\ \rho(-1) & & \\ \boldsymbol{r}_{\mathrm{x}} & \boldsymbol{R}_{[y(2),\boldsymbol{z}(2t)]} & \end{bmatrix}. \tag{35}$$

PLA: Assuming that Eve generates the induced estimate $x(2t+1)$ randomly distributed according to the PDF of the legitimate channel given Eve's observations, i.e., $p_{x(2t+1)|\boldsymbol{z}(2t),\mathcal{H}_0}$ the SAR has been computed in [5]. In particular, let us define $\boldsymbol{S} = \boldsymbol{R}^{-1}_{[x(1)\boldsymbol{z}(2t+1)]}$ and $\boldsymbol{T} = \boldsymbol{R}^{-1}_{[x(2t+1)\boldsymbol{z}(2t+1)]}$ where

$$\begin{aligned} \boldsymbol{R}_{[x(2t+1)\boldsymbol{z}(2t+1)]} &= \mathbb{E}[[x(2t+1), \boldsymbol{z}(2t)]^H [x(2t+1), \boldsymbol{z}(2t)]] \\ &= \begin{bmatrix} 1+\sigma_{\mathrm{x}}^2 & \boldsymbol{r}_{\mathrm{x},2}^H \\ \boldsymbol{r}_{\mathrm{x},2} & \boldsymbol{R}_{\boldsymbol{z}(2t)} \end{bmatrix}, \end{aligned} \tag{36}$$

$$[\boldsymbol{r}_{\mathrm{x},2}]_\ell = [\mathbb{E}[x(2t+1)\boldsymbol{z}^H(2t)]]_\ell = \begin{cases} \alpha_{\mathrm{A}}^* \rho(\ell - 2t + 1) & \ell \text{ even}, \\ \alpha_{\mathrm{B}}^* \rho(\ell - 2t + 1) & \ell \text{ odd}. \end{cases} \tag{37}$$

Partitioning the two matrices as

$$\boldsymbol{S} = \begin{bmatrix} S_{1,1} & \boldsymbol{S}_{1,2} \\ \boldsymbol{S}_{2,1} & \boldsymbol{S}_{2,2} \end{bmatrix} \quad \boldsymbol{T} = \begin{bmatrix} T_{1,1} & \boldsymbol{T}_{1,2} \\ \boldsymbol{T}_{2,1} & \boldsymbol{T}_{2,2} \end{bmatrix} \tag{38}$$

where $T_{1,1}$ and $S_{1,1}$ are scalars, while all other entries are vectors and matrices of suitable dimensions, we define

$$\boldsymbol{E} = \boldsymbol{S}_{2,2} + \boldsymbol{T}_{2,2} - \boldsymbol{R}^{-1}_{\boldsymbol{z}(2t+1)}, \tag{39}$$

and

$$\boldsymbol{V} = \begin{bmatrix} T_{1,1} - \boldsymbol{T}_{1,2}^H \boldsymbol{E}^{-1} \boldsymbol{T}_{1,2} & -\boldsymbol{T}_{1,2}^H \boldsymbol{E}^{-1} \boldsymbol{S}_{1,2} \\ -\boldsymbol{S}_{1,2}^H \boldsymbol{E}^{-1} \boldsymbol{T}_{1,2} & S_{1,1} - \boldsymbol{S}_{1,2}^H \boldsymbol{E}^{-1} \boldsymbol{S}_{1,2} \end{bmatrix}^{-1}. \tag{40}$$

Then (17) becomes

$$R_{\mathrm{PLA}}(2t+1) = \frac{-\ln\det(\boldsymbol{R}_{[x(2t+1),x(1)]}\boldsymbol{V}) + \mathrm{tr}(\boldsymbol{V}\boldsymbol{R}_{[x(2t+1),x(1)]}) - 2}{\ln 2}, \tag{41}$$

where

$$\begin{aligned}\boldsymbol{R}_{[x(2t+1),x(1)]} &= \mathbb{E}[[x(2t+1),x(1)]^H[x(2t+1),x(1)]] \\ &= \begin{bmatrix} 1+\sigma_{\mathrm{x}}^2 & \rho(2t) \\ \rho(-2t) & 1+\sigma_{\mathrm{x}}^2 \end{bmatrix}.\end{aligned} \tag{42}$$

A. Linear Eve's Processing

With linear Eve's processing (LEP) Eve first performs a linear combination of all channel estimates to obtain the MMSE linear estimate of the legitimate channel $\hat{h}$, which is then used for the attack. In particular, for the A-CBCA scheme Eve estimates $h(2)$, while for PLA she estimates $h(1)$. For S-CBCA since we have only bounds on the SAR, Eve can maximize the SAR lower bound. Therefore, from the estimates obtained for PLA and A-CBCA Eve picks the one that minimizes the minimum of the two mutual information, i.e.,

$$\hat{h} = \mathrm{argmin}_{h\in\{\hat{h}(1),\hat{h}(2)\}} \min\{\mathbb{I}(x(1);\boldsymbol{z}(2t)), \mathbb{I}(y(2);\boldsymbol{z}(2t))\}. \tag{43}$$

Let k be the frame index of the desired channel estimate, $k = 1, 2$. We first define the correlation vector

$$\boldsymbol{\beta}(k) = \mathbb{E}[\boldsymbol{z}(2t)h^*(k)] \tag{44}$$

with entries

$$\beta_{2\ell+1}(k) = \alpha_{\mathrm{A}}\rho(k-2\ell-1), \quad \beta_{2\ell}(k) = \alpha_{\mathrm{B}}\rho(k-2\ell). \tag{45}$$

Then we have

$$\boldsymbol{z}(2t) = \boldsymbol{\beta}(k)h(k) + \boldsymbol{w}_0\boldsymbol{R}_{\mathrm{s}}^{1/2}(2t), \tag{46}$$

where $\boldsymbol{w}_0$ is a jointly ZMUPCG $2t$-size vector with i.i.d. entries and the correlation matrix $\boldsymbol{R}_{\mathrm{s}}(2t)$ is

$$\boldsymbol{R}_{\mathrm{s}}(2t) = \mathbb{E}[(\boldsymbol{z}(2t) - \boldsymbol{\beta}(k)h(k))^H(\boldsymbol{z}(2t) - \boldsymbol{\beta}(k)h(k))]. \tag{47}$$

Now, the MMSE estimation of $h(k)$ is obtained as follows

$$\hat{z}(2t) = \frac{1}{\mathrm{tr}(\boldsymbol{R}_{\mathrm{s}}^{-2}(2t))}\boldsymbol{z}(2t)\boldsymbol{R}_{\mathrm{s}}^{-1}(2t)\mathbf{1} = h(k) + \sigma_{\mathrm{z}}(2t)\hat{w}_{\mathrm{z}}(2t)\,, \tag{48}$$

where $\mathbf{1}$ is a $2t$-long column vector of all ones, $\hat{w}_{\mathrm{z}}(2t)$ is ZMUPCG and

$$\sigma_{\mathrm{z}}^2(2t) = \mathrm{tr}^2(\boldsymbol{R}_{\mathrm{s}}^{-2}(2t))\,. \tag{49}$$

In general, LEP is a suboptimal procedure (except for time-invariant channels) in the sense that the attacks by Eve will be less effective. However, this procedure has a limited computational complexity. The SAR obtained with LEP is readily computed from the results of the previous section where $\boldsymbol{z}(2t)$ is replaced by $\hat{z}(2t)$ having unitary correlation with $h(k)$ and noise power $\sigma_{\mathrm{z}}^2(2t)$.

B. LEP with Time-invariant Channels

We now focus on time invariant channels, therefore $h(t) = h$, $q_{\mathrm{A}}(t) = q_{\mathrm{A}}$ and $q_{\mathrm{B}}(t) = q_{\mathrm{B}}$. We also have $\rho(t) = 1\ \forall t$. In this case, the LEP procedure corresponds to optimal processing at Eve.

For $t = 1$ we have

$$\boldsymbol{\beta}(1) = \boldsymbol{\beta}(2) = [\alpha_{\mathrm{A}}, \alpha_{\mathrm{B}}]^T \tag{50}$$

and

$$\boldsymbol{R}_s = \begin{bmatrix} 1 - \alpha_{\mathrm{A}}^2 + \sigma_{\mathrm{v,A}}^2 & 0 \\ 0 & 1 - \alpha_{\mathrm{B}}^2 + \sigma_{\mathrm{v,B}}^2 \end{bmatrix}. \tag{51}$$

For the upper bound of S-CBCA SAR we have

$$\boldsymbol{R}_{[x(1),y(2),\hat{z}(2t)]} = \begin{bmatrix} 1+\sigma_{\mathrm{x}}^2 & 1 & 1 \\ 1 & 1+\sigma_{\mathrm{y}}^2 & 1 \\ 1 & 1 & 1+\sigma_{\mathrm{z}}^2(2t) \end{bmatrix}, \tag{52}$$

$$\begin{aligned} \det \boldsymbol{R}_{[x(1),y(2),\hat{h}(2)]} &= (1+\sigma_{\mathrm{x}}^2)[(1+\sigma_{\mathrm{y}}^2)(1+\sigma_{\mathrm{z}}^2(2)) - 1] - \\ &[(1+\sigma_{\mathrm{z}}^2(2)) - 1] + [1 - (1+\sigma_{\mathrm{y}}^2)] \\ &= \sigma_{\mathrm{x}}^2\sigma_{\mathrm{y}}^2 + \sigma_{\mathrm{x}}^2\sigma_{\mathrm{z}}^2(2) + (1+\sigma_{\mathrm{x}}^2)\sigma_{\mathrm{y}}^2\sigma_{\mathrm{z}}^2(2). \end{aligned} \tag{53}$$

For $t > 1$ note instead that, since $q_{\rm A}$ and $q_{\rm B}$ are the same at all frames, LEP boils down to first estimating

$$\bar{v}_{\rm A}(2t) = \frac{1}{t}\sum_{n=0}^{t-1} v_{\rm A}(2n+1) = \alpha_{\rm A} h + \sqrt{1-\alpha_{\rm A}^2}\, q_{\rm A} + \sigma_{\bar{v},{\rm A}}(2t) w_{{\rm v},{\rm A}}(2t)\,, \tag{54}$$

$$\bar{v}_{\rm B}(2t) = \frac{1}{t}\sum_{n=0}^{t-1} v_{\rm B}(2n) = \alpha_{\rm B} h + \sqrt{1-\alpha_{\rm B}^2}\, q_{\rm B} + \sigma_{\bar{v},{\rm B}}(2t) w_{{\rm v},{\rm B}}(2t)\,, \tag{55}$$

where $w_{{\rm v},{\rm A}}(2t)$ and $w_{{\rm v},{\rm B}}(2t)$ are ZMUPCG and

$$\sigma^2_{\bar{v},{\rm A}}(2t) = \frac{\sigma^2_{{\rm v},{\rm A}}}{t}\,, \quad \sigma^2_{\bar{v},{\rm B}}(2t) = \frac{\sigma^2_{{\rm v},{\rm B}}}{t}\,, \tag{56}$$

and then applying MMSE combining on $\bar{v}_{\rm A}(2t)$ and $\bar{v}_{\rm B}(2t)$ as for the case $t = 1$.

In particular, for $t \to \infty$, from (56) we have that $\sigma^2_{\bar{v},{\rm A}}(2t) = \sigma^2_{\bar{v},{\rm B}}(2t) = 0$, and

$$\hat{z}(2t) = h + \frac{\alpha_{\rm A}(1-\alpha_{\rm A})}{(\alpha_{\rm A}^2+\alpha_{\rm B}^2)} q_{\rm A} + \frac{\alpha_{\rm B}(1-\alpha_{\rm B})}{(\alpha_{\rm A}^2+\alpha_{\rm B}^2)} q_{\rm B}\,, \tag{57}$$

i.e., the Eve's channel estimate is affected only by $q_{\rm A}$ and $q_{\rm B}$.

In Appendix B we derive the SAR with LEP processing of the A-CBCA scheme considering a uniform quantizer with saturation interval $[-v_{\rm sat}, v_{\rm sat}]$ and quantization step Δ.

V. Attacks in the ID-A Phase

In this section we consider attacks by Eve in the ID-A phase for the various UA strategies. Eve transmits together and synchronously with Alice and Bob in the ID-A phase. Therefore she can overlap her signal on the pilots transmitted by Alice. Furthermore, Eve is assumed to be a full-duplex terminal, therefore she can transmit pilots and at the same time receive signals by Alice and Bob thus estimating the channel. Channels are time-invariant (thus we drop index t) and, in order to simplify notation, we assume that Eve has perfect estimates of her channels to both Alice and Bob, i.e., $\sigma^2_{\rm v} = 0$.

Two attacks are considered: pilot contamination (PC) and artificial noise (AN) attack. These attacks are very well known in the literature, and we now apply them to the UA procedures.

a) PC Attack: With PC attack, Eve transmits a scaled version of pilots transmitted by Alice, with scaling factors $\zeta_{\rm A}$ and $\zeta_{\rm B}$, so that Alice and Bob estimate the same channel to Eve, i.e.,

$$\zeta_{\rm A} v_{\rm A} = \zeta_{\rm B} v_{\rm B} = G\,. \tag{58}$$

The estimated channels by Alice and Bob are

$$x(t) = h + G + \sigma_{\rm x} w_{\rm x}(t),\ y(t) = h + G + \sigma_{\rm y} w_{\rm y}(t)\,. \tag{59}$$

In order to analyze the performance of this attack we note that if we divide $x(t)$ and $y(t)$ by $\sqrt{1+|G|^2}$ we obtain again the model (20)-(23) where now $\sigma_{\rm x}^2$ and $\sigma_{\rm y}^2$ become $\sigma_{\rm x}^2/(1+|G|^2)$ and $\sigma_{\rm y}^2/(1+|G|^2)$, respectively. On her side, Eve using LEP obtains the estimate of $(h+G)/\sqrt{1+|G|^2}$

$$\hat{z}_{\rm PC}(2t) = (\hat{z}(2t) + G)/\sqrt{1+|G|^2} \tag{60}$$

with noise variance $\sigma_{\rm z}^2(2t)/(1+|G|^2)$. Therefore the effect of this attack is the scaling of all noise variances, and results of previous Section can be used to compute the SAR. Note that the PC attack has an impact also on PLA correctness, when Eve does not transmit pilots after the ID-A phase. In this case Bob may not recognize the Alice-Bob channel as correct, since G is missing. A-CBCA and S-CBCA are not affected by this issue, since they only use the key extracted in the ID-A phase.

b) AN Attack: With this attack Eve transmits AN during the ID-A phase, i.e. a random ZMUPCG signal aimed at increasing the noise for Alice and Bob. This scenario can be analyzed using the results of the previous section, simply modifying the values of $\sigma_{\rm x}^2$ and $\sigma_{\rm y}^2$. Note that this attack has no impact on the correctness of UA process.

A. Defense Strategies

We describe now possible defense strategies against the ID-A attacks.

- **Random pilots**: legitimate parties use random pilots, locally generated at the transmitter and shared with the legitimate receiver after the ID-A phase on the public authenticated error-free channel (which as we have seen, must be in any case available in the ID-A phase). In this case Eve would not be able to add coherently her pilot and induce a desired channel;
- **Channel and noise power estimation**: if reference values of these powers are available at the legitimate receivers, the attack can be detected (see also [19]);

- **Channel agreement**: as outlined in [19] by estimating the channel at both Alice and Bob and comparing the estimates without disclosing them to Eve it is possible to check if the two legitimate users see the same channel, thus preventing Eve from performing an attack in which she does not know the channels to Alice and Bob.

Moreover, note that the described attacks require the knowledge of the Alice-Eve and Bob-Eve channels before transmissions, therefore implementing the ID-A stage at the very beginning of transmission would prevent Eve from getting the channel estimates and deploy the attack.

VI. Numerical Results

We provide now some results on the SAR of the various UA systems. We focus in particular on the Rayleigh AWGN reciprocal channels of Section IV, where both Alice and Bob transmit with unitary power and channels are vectors of i.i.d. ZMUPCG. We consider both time-invariant and time-variant channels, and ID-A attacks described in Section V. For A-CBCA we have already observed that the SAR can be made arbitrarily large in the presence of a passive eavesdropper: here we report the results for a uniform quantizer with 3 bits (corresponding to 8 quantization levels) and saturation value v_{sat} that ensures a probability of saturation of 10^{-2}. Unless differently specified we consider $\sigma_{\mathrm{x}}^2 = \sigma_{\mathrm{y}}^2 = \sigma_{\mathrm{v,A/B}}^2 = -10$ dB.

A. Time-invariant Channels

We start from time-invariant channels. Fig. 1 shows the SAR versus (vs) the correlation coefficients $\alpha_{\mathrm{A}} = \alpha_{\mathrm{B}}$, at ID-V frame $t = 3$, i.e., immediately after the two ID-A frames. The results for the lower and upper bound on the SAR of S-CBCA are shown as $R_{\mathrm{S-CBCA}}$ low and up, respectively. As expected, with an increasing correlation among the legitimate and eavesdropper's channels the SAR decreases. Note that even when $\alpha_{\mathrm{A}} = \alpha_{\mathrm{B}} = 1$ we still may have a non-zero SAR. In particular, in A-CBCA Bob benefits from the randomness of the noise which is assumed independent with respect to that of Eve. Similarly, in PLA Eve generates a random attack channel having as mean her channel estimate rather than the estimate obtained by the legitimate user in the ID-A steps, and the two estimates differ due to the noise. The lower bound for S-CBCA is indeed zero in this case.

Fig. 2 shows the SAR vs the ID-V frame for three values of channel correlations: $\alpha_{\mathrm{A}} = \alpha_{\mathrm{B}} = 0.1$ (solid lines), $\alpha_{\mathrm{A}} = \alpha_{\mathrm{B}} = 0.4$ (dashed lines), and $\alpha_{\mathrm{A}} = \alpha_{\mathrm{B}} = 0.8$ (dotted lines). We observe that for all schemes, as the ID-V frame t increases SAR decreases, because in the meantime Eve

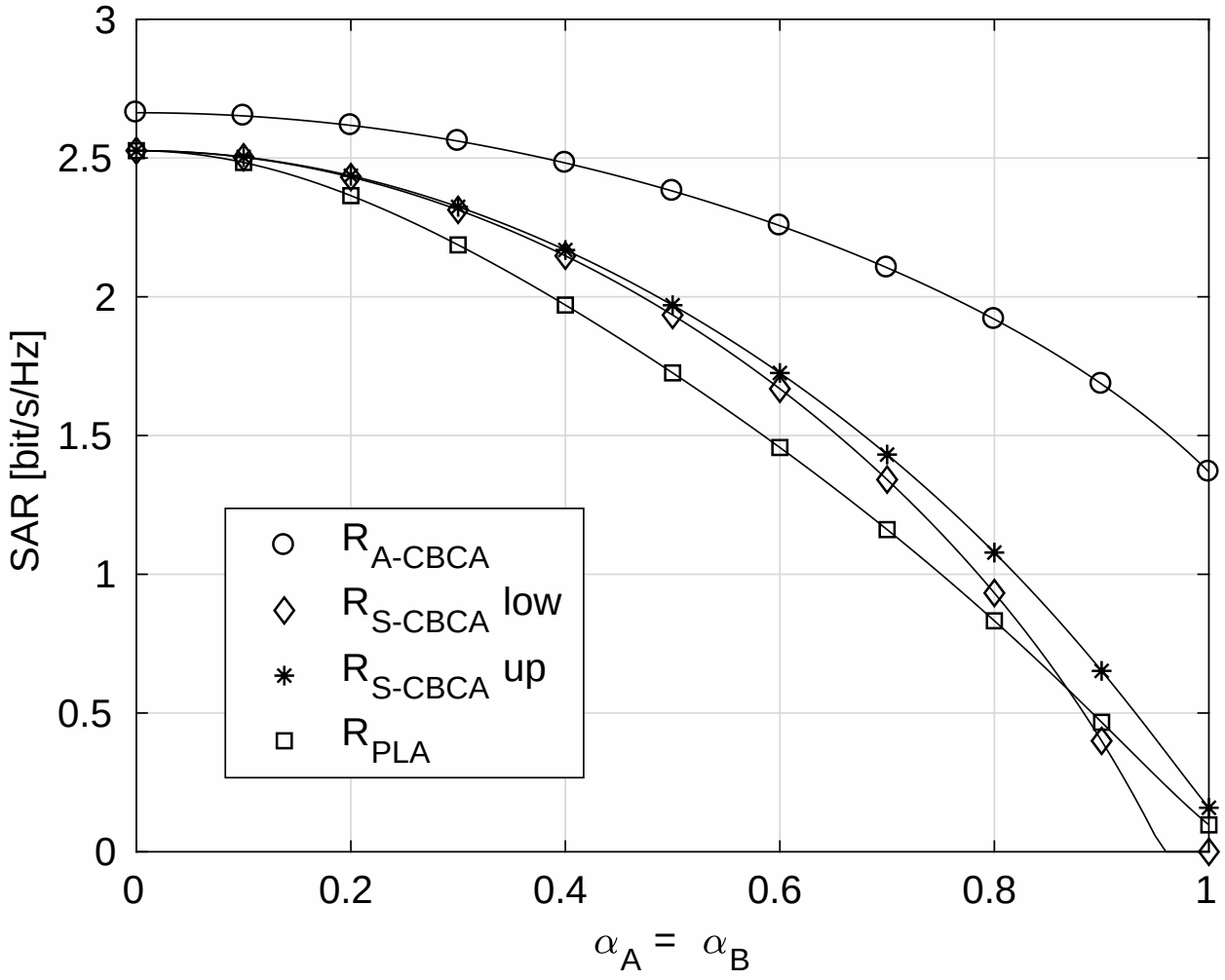

Figure 1. SAR vs the correlation coefficients $\alpha_{\rm A} = \alpha_{\rm B}$ for time-invariant channels at ID-V frame $t = 3$.

has obtained a better channel estimate. The degradation of authentication performance is more remarkable for a higher value of channel correlation factor, since in this case having a more accurate knowledge of her channels to Alice and Bob truly provides Eve a better knowledge of the Alice-Bob channel.

B. Time-varying Channels

We consider now frame-time-variant channels with Jakes fading. In particular the channel is time-invariant in each frame while the evolution over frames is

$$h(t) = \rho(t)h(1) + \sqrt{1 - |\rho(t)|^2}g(t)\,, \tag{61}$$

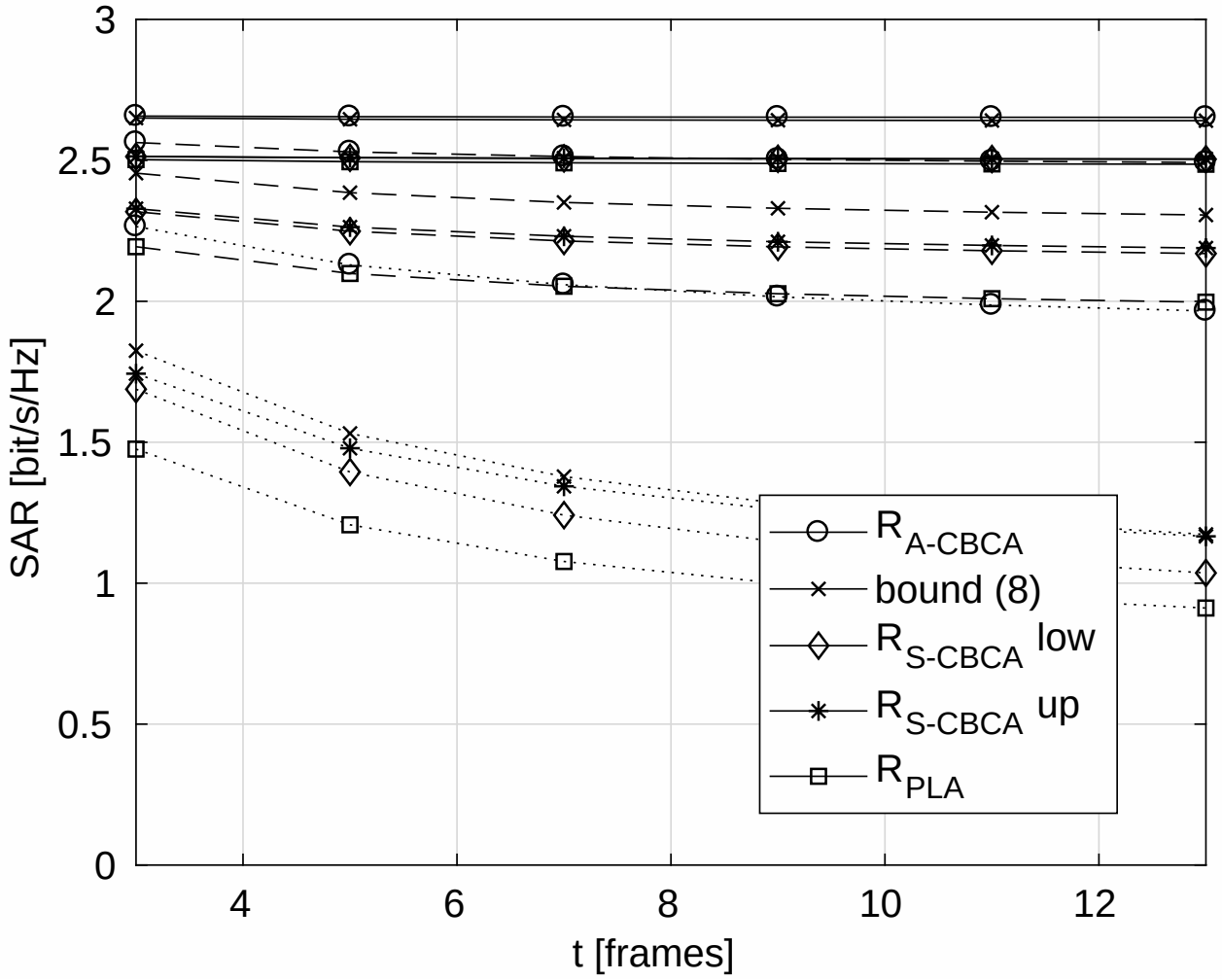

Figure 2. SAR vs ID-V frame index, for three values of channel correlations: $\alpha_{\mathrm{A}} = \alpha_{\mathrm{B}} = 0.1$ (solid lines), $\alpha_{\mathrm{A}} = \alpha_{\mathrm{B}} = 0.4$ (dashed lines), and $\alpha_{\mathrm{A}} = \alpha_{\mathrm{B}} = 0.8$ (dotted lines).

with $g(t)$ ZMUPCG and

$$\rho(t) = J_0(2\pi f_d t T)\,, \tag{62}$$

with T the frame duration, f_d the Doppler frequency and $J_0(\cdot)$ the zero-order Bessel function of the first kind.

As we have observed, for PLA channel variations have an impact on the ability of the scheme to effectively authenticate a legitimate transmission, since the channel (which is used as user signature) changes. Therefore Fig. 3 shows the SAR vs the ID-V frame index, for three values of the normalized Doppler frequency $T f_{\mathrm{D}}$ for the PLA scheme when $\alpha_{\mathrm{A}} = \alpha_{\mathrm{B}} = 0.4$. We assume that Eve estimates the channel in only the first two frames, and up to frame $t > 2$ Alice and

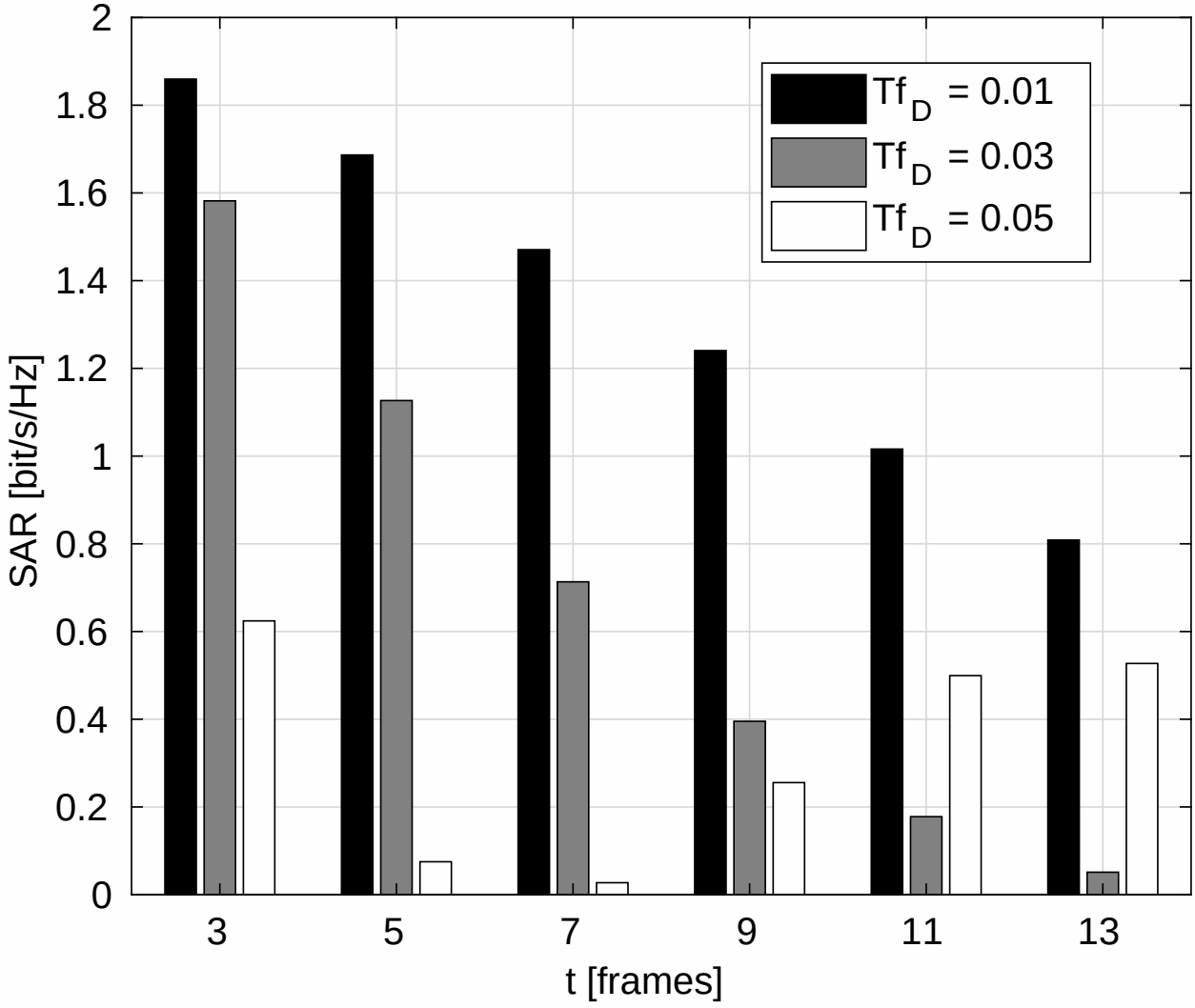

Figure 3. SAR vs the ID-V frame index, for three values of the normalized Doppler frequency $Tf_{\rm D}$ for the PLA scheme. $\alpha_{\rm A} = \alpha_{\rm B} = 0.4$.

Bob are not transmitting. We observe that as the normalized Doppler frequency increases the SAR decreases. Moreover, as ID-V frame index increases the SAR is reduced as well. In both cases the channel variations prevent an effective authentication. Lastly, note that for the highest value of the normalized Doppler frequency the SAR increases for a higher number of frames: this is due to the Jakes model, and in particular to the behavior of the correlation (62) as f_d first decreases and then increases.

We now consider the effect of time-variations on all the schemes. Fig. 4 shows the SAR as a function of the normalized Doppler frequency for three ID-V frames $t = 3$ (solid lines), $t = 5$ (dashed lines) and $t = 9$ (dotted lines). Also in this case $\alpha_{\rm A} = \alpha_{\rm B} = 0.4$. We observe that the

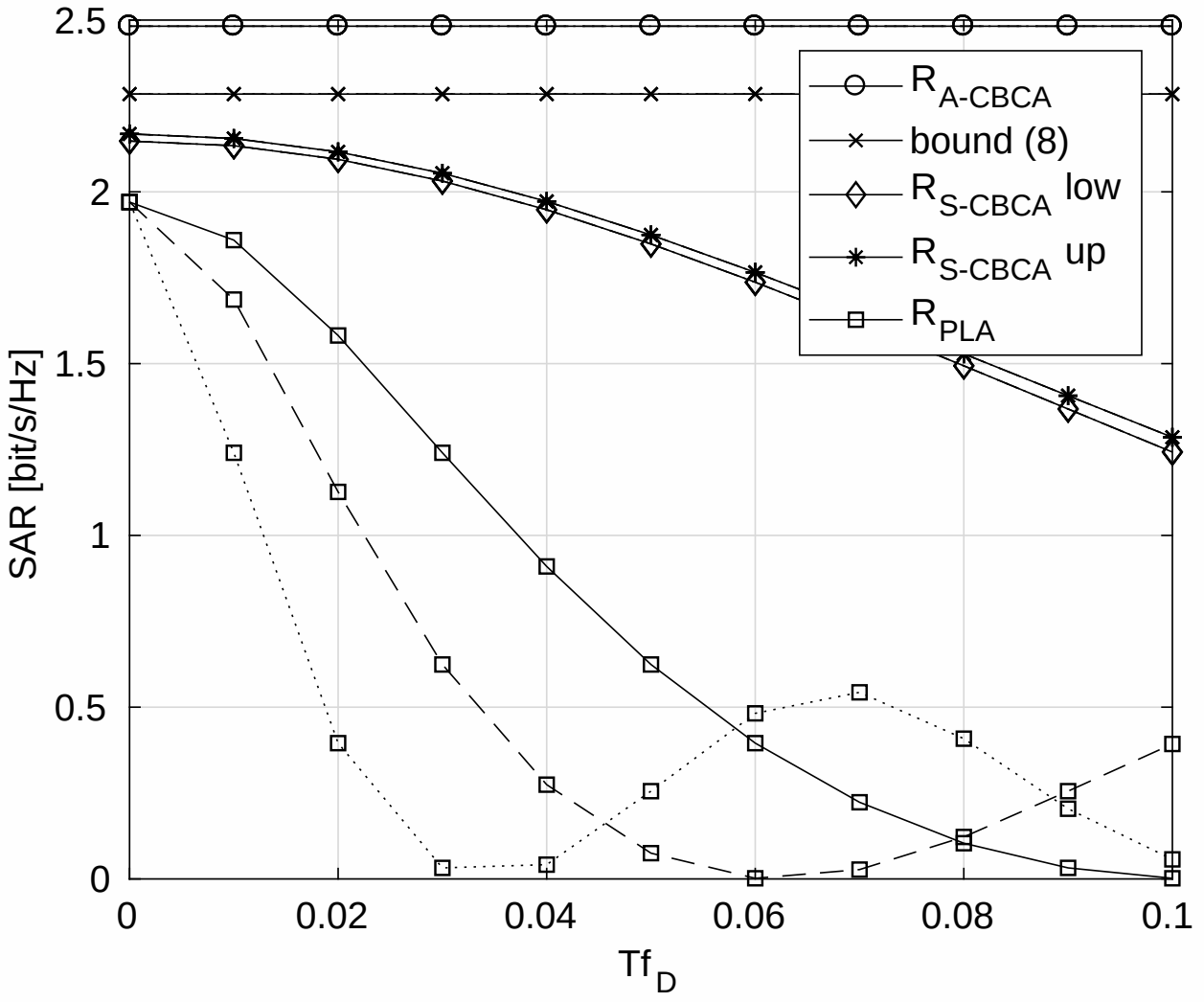

Figure 4. SAR vs the normalized Doppler frequency for three ID-V frames $t = 3$ (solid lines), $t = 5$ (dashed lines) and $t = 9$ (dotted lines). $\alpha_{\mathrm{A}} = \alpha_{\mathrm{B}} = 0.4$.

A-CBCA scheme is not affected by the time-variation of the channel across frames, as it only uses one frame. Moreover, for S-CBCA the SAR decreases for increasing normalized Doppler frequency but it is insensitive to the frame in which authentication is performed: in fact, for this scheme the channel is used only in the ID-A phase to establish the secret key and channel variations in further frames are not relevant. On the other hand, S-CBCA and PLA are more heavily affected, since channel variations have an impact on both ID-A and ID-V phases. Also in this case we observe the effect of increasing channel correlation for high Doppler frequency, as already observed for Fig. 2.

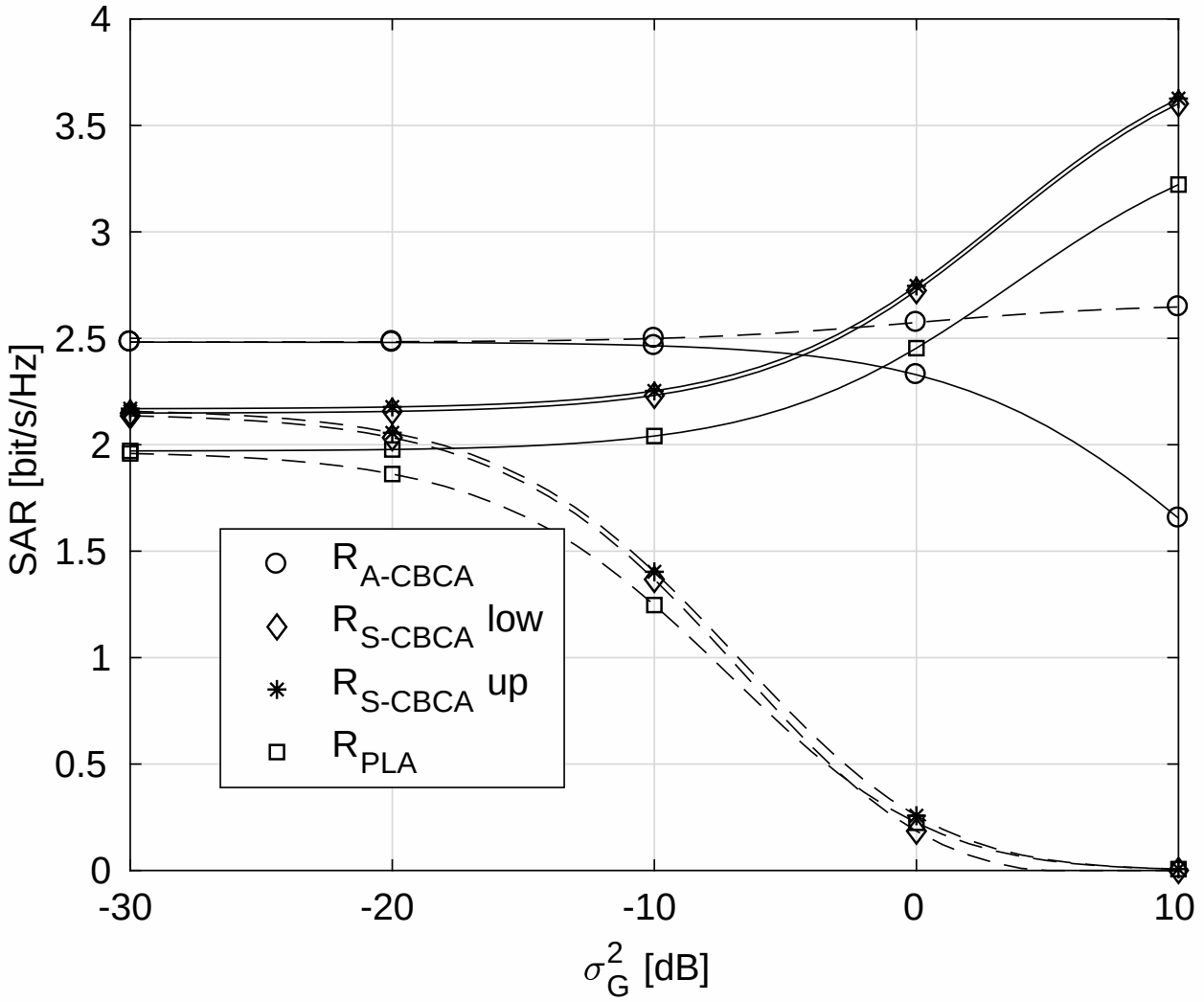

Figure 5. SAR as a function of σ_G^2 for ID-A-phase attacks, for both the PC attack (solid lines) and the AN attack (dashed lines).

C. Active Attacks

We now consider the active attacks by Eve, as described in Section V, namely PC and AN attack. Fig. 5 shows the SAR as a function of σ_G^2 for PC attack (solid lines) and AN attack (dashed lines). Also in this case $\alpha_{\mathrm{A}} = \alpha_{\mathrm{B}} = 0.4$ and channels are time-invariant. We recall that for the PC attack σ_{G}^2 is the power of the random channel superimposed to the effective channel that is estimated by the legitimate users in the ID-A phase. From the figure we observe that this attack is not effective for S-CBCA and PLA schemes, instead increasing the SAR. This is due to the fact that this attack is equivalent to a reduction of the noise estimate for both legitimate users and Eve, thus resulting in a globally favorable situation. For A-CBCA this attack

in instead effective, since the reduction of the uncertainty on the channel estimated by Bob helps Eve to perform a more effective attack. When we consider the AN attack instead we observe that it is effective for the S-CBCA and PLA schemes, while it yields a higher SAR (thus not being effective) for A-CBCA. In fact, since the AN is not used by Eve in its estimation phase, this attack provides additional randomness to Alice, that is unknown to Eve, thus supporting the extraction of random bits from the channel in A-CBCA. Instead, for the other schemes the additional noise reduces the capabilities of extracting a shared secret key from the channel in the S-CBCA scheme, or the possibility of effectively detecting channel variations in PLA scheme. Therefore this attack is effective against these two schemes. As we mentioned, randomizing pilots and channel agreement can be effective in preventing these attacks.

VII. Conclusions

In this paper we have compared three UA strategies, based on either symmetric/asymmetric keys or on physical layer authentication, where in all cases the features used for authenticating the user are extracted from the communication channel. The comparison has been performed in terms of SAR, i.e., the rate at which the probability that UA security is broken goes to zero as the number of used channel resources goes to infinity. From the analysis and the numerical results we can conclude that the A-CBCA scheme provides potentially the highest SAR and is immune to channel changes. Then, the S-CBCA scheme, which uses the source-based SKA, is slightly more sensitive to the channel variations but has a lower bound on SAR that typically is higher than the SAR achieved with PLA. Moreover, PLA is the most sensitive solution to channel variations.

References

[1] W. Stallings, *Cryptography and Network Security: Principles and Practice*, 5th ed. Upper Saddle River, NJ, USA: Prentice Hall Press, 2010.

[2] G. J. Simmons, "Authentication theory/coding theory," in *Proc. Advances in Cryptology—CRYPTO 84 (Lecture Notes in Computer Science)*, G. Blakley and D. Chaum, Eds. Berlin, Germany: Springer-Verlag, 1985, pp. 411–431.

[3] G. Simmons, "A survey of information authentication," vol. 76, no. 5, pp. 603–620, May 1988.

[4] 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, Oct 2015.

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

[6] T. Daniels, M. Mina, and S. F. Russell, "Short paper: A signal fingerprinting paradigm for general physical layer and sensor network security and assurance," in *Proc. First Int. Conf. on Security and Privacy for Emerging Areas in Comm. Networks (SECURECOMM'05)*, Sept 2005, pp. 219–221.

[7] L. Xiao, L. J. Greenstein, N. Mandayam, and W. Trappe, "Channel-based spoofing detection in frequency-selective Rayleigh channels," *IEEE Trans. Wireless Comm.*, vol. 8, no. 12, pp. 5948–5956, Dec. 2009.

[8] S. Jiang, "Keyless authentication in a noisy model," *IEEE Trans. on Information Forensics and Security*, vol. 9, no. 6, pp. 1024–1033, June 2014.

[9] J. Liu and X. Wang, "Physical layer authentication enhancement using two-dimensional channel quantization," *IEEE Trans. on Wireless Comm.*, vol. 15, no. 6, pp. 4171–4182, June 2016.

[10] L. Xiao, T. Chen, G. Han, W. Zhuang, and L. Sun, "Game theoretic study on channel-based authentication in MIMO systems," *IEEE Trans. on Vehicular Technology*, 2017, IEEE Early Access Articles.

[11] J. Zhang, T. Q. Duong, A. Marshall, and R. Woods, "Key generation from wireless channels: A review," *IEEE Access*, vol. 4, pp. 614–626, 2016.

[12] S. Jiang, "On the optimality of keyless authentication in a noisy model," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 6, pp. 1250–1261, June 2015.

[13] R. Ahlswede and I. Csiszar, "Common randomness in information theory and cryptography. i. secret sharing," *IEEE Trans. on Information Theory*, vol. 39, no. 4, pp. 1121–1132, Jul 1993.

[14] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. on Information Theory*, vol. 39, no. 3, pp. 733–742, May 1993.

[15] S. Tomasin, "Comparison between asymmetric and symmetric channel-based authentication for MIMO systems," in *Proc. 21st Int. ITG Workshop on Smart Antennas*, Mar. 2017.

[16] M. Bloch and J. Barros, *Physical-Layer Security. From Information Theory to Security Engineering*. Cambridge: Cambridge University Press, 2011.

[17] E. T. Jaynes, "Information theory and statistical mechanics," *K. Ford. Statistical Physics*, p. 181, 1963.

[18] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, 2006.

[19] S. Tomasin, I. Land, and F. Gabry, "Pilot contamination attack detection by key-confirmation in secure MIMO systems," in *Proc. IEEE Global Conf. Commun. (GLOBECOM)*, 2016.

[20] 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 Trans. on Inf. Forensics and Security*, vol. 10, no. 5, pp. 941–952, May 2015.

APPENDIX A
PROOF OF LEMMA 2

We first observe that by the chain rule

$$\begin{aligned}&\mathbb{D}\big(p_{x(1),x(2t+1)|\mathcal{H}_0}\big|\big|p_{x(1),x(2t+1)|\mathcal{H}_1}\big) \leq \\ &\mathbb{D}\big(p_{x(1),x(2t+1),\boldsymbol{z}(2t)|\mathcal{H}_0}\big|\big|p_{x(1),x(2t+1),\boldsymbol{z}(2t)|\mathcal{H}_1}\big)\,.\end{aligned} \tag{63}$$

Moreover, the KL divergence of the joint PDFs can be written as the expectation of the conditioned PDFs, i.e.,

$$\begin{aligned}\mathbb{D}(p_{x(1),x(2t+1),\boldsymbol{z}(2t)|\mathcal{H}_0}||p_{x(1),x(2t+1),\boldsymbol{z}(2t)|\mathcal{H}_1}) = \\ \mathbb{E}[\mathbb{D}(p_{x(1),x(2t+1)|\mathcal{H}_0,\boldsymbol{z}(2t)}||p_{x(1),x(2t+1)|\mathcal{H}_1,\boldsymbol{z}(2t)})]\,,\end{aligned} \tag{64}$$

where expectation is taken with respect to $\boldsymbol{z}(2t)$. Now, we also have the general relation between mutual information and KL divergence for random variables a, b and c

$$\mathbb{I}(a;b|c) = \mathbb{E}_c[\mathbb{D}(p_{a,b|c}||p_{a|c}p_{b|c})]. \tag{65}$$

Recalling that conditionally on $\boldsymbol{z}(2t)$, $x(1)$ and $x(2t+1)$ (as generated by Eve) are independent [20] we have

$$p_{x(1),x(2t+1)|\mathcal{H}_1,\boldsymbol{z}(t)} = p_{x(1)|\mathcal{H}_1,\boldsymbol{z}(t)}p_{x(2t+1)|\mathcal{H}_1,\boldsymbol{z}(2t)}\,. \tag{66}$$

Now, assuming that Eve does not attack in the ID-A frames, $p_{x(1)|\mathcal{H}_1,\boldsymbol{z}(t)} = p_{x(1)|\mathcal{H}_0,\boldsymbol{z}(t)}$. Moreover, if the attack has PDF $p_{x(2t+1)|\mathcal{H}_0,\boldsymbol{z}(2t)}$, we have

$$p_{x(2t+1)|\mathcal{H}_1,\boldsymbol{z}(2t)} = p_{x(2t+1)|\mathcal{H}_0,\boldsymbol{z}(2t)} \tag{67}$$

and therefore

$$p_{x(1),x(2t+1)|\mathcal{H}_1,\boldsymbol{z}(t)} = p_{x(1)|\mathcal{H}_0,\boldsymbol{z}(t)}p_{x(2t+1)|\mathcal{H}_0,\boldsymbol{z}(2t)}\,, \tag{68}$$

and using (63) we have

$$\begin{aligned}\mathbb{D}(p_{x(1),x(2t+1)|\mathcal{H}_0}||p_{x(1),x(2t+1)|\mathcal{H}_1}) \leq \\ \mathbb{I}(x(1);x_{\mathcal{H}_0}(2t+1)|\boldsymbol{z}(2t))\,,\end{aligned} \tag{69}$$

which provides (19).

APPENDIX B
SAR FOR A-CBCA WITH RAYLEIGH FADING

We first observe that the Rayleigh channel coefficient h has i.i.d. real and imaginary parts, therefore the SAR will be twice of the rate obtained considering the quantization of the real

part, i.e.,

$$R_{\mathrm{A-CBCA}}(2t+1) = -2\sum_{i=0}^{M-1}\int p_{\langle y(2)\rangle|\hat{z}(2t)}(a_i|b)p_{\hat{z}(2t)}(b) \log p_{\langle y(2)\rangle|\hat{z}(2t)}(a_i|b)\mathrm{d}b\,, \tag{70}$$

For time-invariant channels and considering (T_{i-1}, T_i) as quantization interval for a_i we have

$$\begin{aligned} &p_{\langle y(2)\rangle|\hat{z}(2t)}(i|b) = \frac{1}{p_{\hat{z}(2t)}(b)}\times \\ &\mathbb{P}\left(\hat{h}+\frac{\sigma_{\mathrm{y}}}{\sqrt{2}}w_{\mathrm{y}}(2)\in(T_{i-1},T_i], \hat{h}+\frac{\sigma_{\mathrm{z}}(2t)}{\sqrt{2}}\hat{w}_{\mathrm{z}}(t)=b, h=\hat{h}\right) \\ &=\int_{-\infty}^{\infty}\mathbb{P}\left(h+\frac{\sigma_{\mathrm{y}}}{\sqrt{2}}w_{\mathrm{y}}(2)\in(T_{i-1},T_i]\right)\times \\ &\quad p_{\hat{w}_{\mathrm{z}}(2t)}\left(-\frac{h\sqrt{2}}{\sigma_{\mathrm{z}}(2t)}+\frac{b\sqrt{2}}{\sigma_{\mathrm{z}}(2t)}\right)p_h(h)dh\,, \end{aligned} \tag{71}$$

where we used half of all noise variances since we are considering only the real part of the channel estimates. Now considering v_{sat} as the saturation value and $T_i = v_{\mathrm{sat}} + \Delta i$, $T_{-1} = -\infty$, $T_M = \infty$, we have that the first function in (70) is

$$\begin{aligned} &\mathbb{P}\left(h+\frac{\sigma_{\mathrm{y}}}{\sqrt{2}}w_{\mathrm{y}}(2)\in(T_{i-1},T_i]\right) = \\ &\begin{cases} 1-\mathrm{Q}\left(\frac{\sqrt{2}(-v_{\mathrm{sat}}+\Delta-h)}{\sigma_{\mathrm{y}}}\right) & i=0 \\ \mathrm{Q}\left(\frac{\sqrt{2}(-v_{\mathrm{sat}}+\Delta i-h)}{\sigma_{\mathrm{y}}}\right)-\mathrm{Q}\left(\frac{\sqrt{2}(-v_{\mathrm{sat}}+\Delta(i+1-h)}{\sigma_{\mathrm{y}}}\right) & i\in[1,M-2] \\ \mathrm{Q}\left(\frac{\sqrt{2}(-v_{\mathrm{sat}}+\Delta(M-1)-h)}{\sigma_{\mathrm{y}}}\right) & i=M-1\,, \end{cases} \end{aligned} \tag{72}$$

where $\langle y(2)\rangle$ is the quantization index of $y(2)$. Moreover for second and third functions in (70) we have

$$p_h(\hat{h}) = \frac{1}{\pi}e^{-\hat{h}^2}\,, \tag{73}$$

$$p_{\hat{w}_{\mathrm{z}}(2t)}\left(-\frac{\sqrt{2}h}{\sigma_{\mathrm{z}}(2t)}+\frac{\sqrt{2}b}{\sigma_{\mathrm{z}}(2t)}\right) = \frac{1}{\sigma_{\mathrm{z}}(2t)\sqrt{\pi}}e^{-\frac{(h-b)^2}{\sigma_{\hat{z}}^2(2t)}}\,. \tag{74}$$

Both integrals in (71) and in (70) must be solved by numerical methods.