Physical Layer Security of Massive MIMO Spatially-uncorrelated Rician Channels

Giang Quynh Le Vu*, Nhat Thang Le†, and Kien Trung Truong‡
* Faculty of Information Technology, National Academy of Education Management, Hanoi, Vietnam
† Posts and Telecommunication Institute of Technology, Hanoi, Vietnam
‡ Undergraduate Faculty, Fulbright University Vietnam, Ho Chi Minh City, Vietnam.
Email: quynhgiang81@gmail.com, thangln@ptit.edu.vn, kien.truong@fulbright.edu.vn

***Abstract*—Pilot contamination jamming not only reduces the secrecy capacity but also is difficult to detect. In this paper, we proposed a technique, in which we employ random N-PSK to detect passive eavesdropper throw pilot contamination in Massive MIMO Uncorrelated Rician Fading Channels. The technique only needs two training times slot and without channel knowledge.**

***Index Terms*—Massive MIMO, physical layer security, Rician fading, eavesdropper detection, jammer detection.**

I. Introduction

Massive Multiple-Input Multiple-Output (MIMO) is one of the key transmission techniques in the 5th Generation (5G) New Radio (NR) networks as well as in the 6th Generatino (6G) networks [1], [2], [3], [4]. In such systems, the base station (BS) with a larger number of antennas serve one or more single-antenna users. Previous work showed that under certain conditions, if the number of antennas at the base station is large enough, the channels between the BS and the users are orthogonal to each other, making the effects of noise and co-channel interference negligible. Notably, this orthogonality of the transmission channels makes massive MIMO systems in Rayleigh fading highly secure at the physical layer [5], [6].

Because of the characteristics of the radio environment, un-legitimated devices can affect security, integrity, and availability of information by two ways: i) passive eavesdropper [7], [8], [9] , ii) active attacker or jammer) [5], [10], [11]. Specifically, passive eavesdroppers listen and try to decode transmission's signal from the transmitter. This impact reduces the performance of the legitimated communication systems, even interrupt system. This paper focused on detection passive eavesdropper in Massive MIMO uncorrelated Rician fading channel. Noted that, almost previous studies relevant to systems, which has one passive eavesdropper assumed the Rayleigh fading channels. Theoretically, the Rician fading channels this paper is more general than the Rayleigh fading channels, because it included Line-Of-Sight component [12]. In [13], [14] provided the analysis of the secrecy capacity analysis of the point-to-point transmission system with a finite number of antennas at the base station. In particular, our analytical results show that under the Rician fading channels, the eavesdropping data rate is increased with the number of antennas at the base station while the secrecy capacity of the system approaches a saturation values as the antennas number of base station reach infinity. In this paper, we propose a scheme to detect the presence of Eve who attacks on the channel estimation using the method in [7]. The main idea is to make use of random pilots for channel estimation. We use phase-shift keying (N-PSK) symbols as the pilot symbols which are transmitted randomly. The scalar product between the received vectors has been developed detection the presence of Eve.

This paper is organized as: Section II, we introduce system model. Section III we present detection procedure based on random training pilots in the presence of received noise and construction of the detection regions. Section IV there are simulation results and Section V concludes the paper.

Notation: a is scalar, $\mathbf{a}$ is vector, $\mathbf{A}$ is matrix, $[\mathbf{A}]_{i,j}$represents (i,j), $\mathbf{A}^H$, is Hermitian matrix transposed, $\mathbb{E}[\dot{]}$ is main value.

II. System Model

Consider a single-cell massive multiple-input multiple-output (MIMO) system where a base station(BS) A communicates legitimate user B under the presence of a passive eavesdropper E both that has a single antenna. For notations convenience, denote $\mathcal{X} = \{\mathrm{B}, \mathrm{E}\}$ in Fig 1. Although this is a system model that has been simplified by considering only two users, the results of this model can be easily extended to cases where there are multiple users. BS equipped M antennas, while $M \gg 2$. For convenience, we're considering that antenna array of base station A is Uniform Linear Array (ULA). Noted d_{BS} is the distance between the adjacent antennas at the BS, $\bar{d}_{\mathrm{BS}} = 2\pi d_{\mathrm{BS}}/\lambda$ is the distance between adjacent antennas at the base base station normalized with the number of waves, where λ is the wavelength

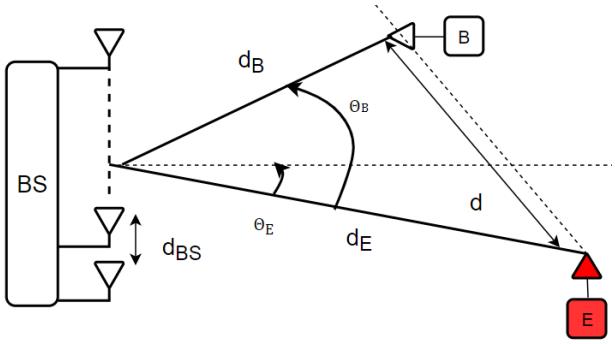

Fig. 1. System model, in which BS communicates with Bob (B) and eavesdropper (E).

corresponding to the carrier frequency, d_k, $\forall k \in \mathcal{K}$, is the distance from the base station to UE$_k$ [B,E]. θ_k, $\forall k \in \mathcal{K}$, is the angle created by the beam connected from the BS to UE$_k$ [B,E] and boresight. The value range of θ_k is from $-\pi$ to π. Fig 1 show the one in which scenarios. Assume frequency-flat block-fading channel model where channel coefficients keep unchanged during the duration of each radio frame of τ symbols and change independently frame-by-frame. Let $\mathbf{h}_{\mathrm{B}} \in \mathbb{C}^{N_{\mathrm{t}} \times 1}$ be the channel coefficient vector between legitimate user B and base station A. Let $\mathbf{h}_{\mathrm{E}} \in \mathbb{C}^{N_{\mathrm{t}} \times 1}$ be the channel coefficient vector between eavesdropper E and base station A. In this paper, we consider Rician fadig channel model. In particular, each channel coefficient vector $\mathbf{h}_{\mathrm{X}}$, where $\mathrm{X} \in \mathcal{X}$, is modeled as a realization of the circularly symmetric complex Gaussian distribution

A. LOS Propagation Model

In this section, we conduct LoS propagation model of system vector channel $\mathbf{g}_k, \forall k \in \mathcal{K}$. Noted $\mathbf{a}_{\mathrm{BS}}(\theta) \in \mathbb{C}^M \times 1$ is vector of M antennas at BS at the direction corresponding to the wave propagation angle θ. Noted $\bar{\mathbf{g}}_k \in \mathbb{C}^M \times 1$ is LOS propagation $\mathbf{g}_k$ with $k \in \mathcal{K}$. When, $\bar{\mathbf{g}}_k$ was determined.

$$\bar{\mathbf{g}}_k = \frac{1}{\sqrt{M}} \left[1 \; e^{j\bar{d}_{\mathrm{BS}} \sin\theta} \; \cdots \; e^{j\bar{d}_{\mathrm{BS}}(M-1)\sin\theta}\right]^T. \tag{1}$$

B. NLOS Propagation Model

Noted $\mathbf{h}_k \in \mathbb{C}^{N_{\mathrm{t}} \times 1}$ is NLOS propagation coefficient of $\mathbf{h}_k$, $\forall k \in \mathcal{K}$. With the rich scatter wave propagation, we can assume that $\mathbf{h}_k \sim \mathcal{CN}(\mathbf{0}_{M\times 1}, \mathbf{R}_k)$, while $\mathbf{R}_k = \mathbb{E}[\mathbf{h}_k \mathbf{h}_k^H] \in \mathbb{C}^M \times M$ is the spatial correlation matrix of the NLOS coefficient transmissions vector $\mathbf{h}_k$ where $\mathrm{tr}(\mathbf{R}_k) = 1$. The spatial correlation matrix $\mathbf{R}_k$ can be determined by the following model: Scattering cluster channel model:
Noted θ_k and σ_θ^2 is main value and variance of angles corresponds to the transmitted vector ,$p_\theta(\phi)$ is the probability distribution density function of the angles corresponding to the wave propagation rays being scattered across the same cluster. Then the rows element m rows column n of $\mathbf{R}_k$ are $\forall m, n \in \{1, 2, \cdots, M\}$

$$[\mathbf{R}_k]_{m,n} = \xi \int_{-\pi}^{\pi} e^{j\bar{d}_{\mathrm{BS}} \sin(\theta_k + \phi)} p_\theta(\phi) d\phi. \tag{2}$$

Closed form of $[\mathbf{R}_k]_{m,n}$, follow the truncated Laplacian. For example, if θ is truncated Laplacian distribution then $[\mathbf{R}_k]_{m,n}$ approximately equal (when σ_θ less 10^o) [15]

$$[\mathbf{R}_k]_{m,n} \approx \frac{\xi e^{j\bar{d}_{\mathrm{BS}}|m-n|\sin\theta_k}}{1 + \frac{\sigma_\theta^2}{2}[d_{\mathrm{BS}}(m-n)\cos\theta_j]^2}. \tag{3}$$

From the above assumptions we build up the uplink channel model of system. Noted $\mathbf{h}_{\mathrm{B}} \in \mathbb{C}^{M\times 1}$ is the channel vector on uplink from the user to the base station and $\mathbf{h}_{\mathrm{E}} \in \mathbb{C}^{M\times 1}$ is channel vector on uplink from eavesdropper to base station. Consider channel model is perfect TDD reciprocity $\mathbf{h}_{\mathrm{B}}^H, \mathbf{h}_{\mathrm{E}}^H \in \mathbb{C}^{1\times M}$. In this paper, we suppose κ_X coefficient Rician and β_{X} is large-scale fading of channel from base station to destination $\mathrm{X} \in \mathcal{X}$. large-scale fading coefficient correspond part (LOS: Line-of-Sight)$\beta_{\mathrm{X,L}}$ and (NLOS: Non Line-of-Sight) $\beta_{\mathrm{X,N}}$ given by

$$\beta_{\mathrm{X,L}} = \sqrt{\frac{\kappa_{\mathrm{X}}}{\kappa_{\mathrm{X}}+1}}\beta_{\mathrm{X}}; \; \beta_{\mathrm{X,N}} = \sqrt{\frac{1}{\kappa_{\mathrm{X}}+1}}\beta_{\mathrm{X}}. \tag{4}$$

Then channel vector $\mathbf{h}_{\mathrm{X}}$ has the distribution $\mathbf{h}_{\mathrm{X}} \sim \mathcal{CN}(\mathbf{g}_{\mathrm{X}}, \beta_{\mathrm{X,N}}\mathbf{I}_N)$ for $\mathrm{X} \in \mathcal{X}$ given by

$$\mathbf{h}_{\mathrm{X}} = \mathbf{g}_{\mathrm{X}} + \sqrt{\beta_{\mathrm{X,N}}}\mathbf{w}_{\mathrm{X}}. \tag{5}$$

The $\mathbf{g}_{\mathrm{X}}$ is LOS coefficient and $\mathbf{w}_{\mathrm{X}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_{\mathrm{t}}})$ is small-scale fading coefficient. Then, channel vector coefficient corresponds LOS from BS to destination $\mathrm{X} \in \mathcal{X}$ given by

$$\mathbf{g}_{\mathrm{X}} = \beta_{\mathrm{X,L}}^{1/2}\left[1 \; e^{j2\pi d\sin\theta_{\mathrm{X}}} \; \cdots \; e^{j2\pi d(M-1)\sin\theta_{\mathrm{X}}}\right]^T. \tag{6}$$

in that θ_{X} is angle of arrival departure from X to BS and d is function of the antenna spacing and wave length . Note that $\mathbf{g}_{\mathrm{X}}^H \mathbf{g}_{\mathrm{X}} = M\beta_{X,L}$ and $\forall X \in \mathcal{X}$. It is convenient to denote

$$\psi(\theta_{\mathrm{B}}, \theta_{\mathrm{E}}) = \pi d_{\mathrm{BE}}(\sin\theta_{\mathrm{B}} - \sin\theta_{\mathrm{E}}) \tag{7}$$

$$\alpha(\theta_{\mathrm{B}}, \theta_{\mathrm{E}}, M) = \frac{\sin(M\psi(\theta_{\mathrm{B}}, \theta_{\mathrm{E}}))}{\sin(\psi(\theta_{\mathrm{B}}, \theta_{\mathrm{E}}))}. \tag{8}$$

By some computing, we have

$$\mathbf{g}_{\mathrm{E}}^H \mathbf{g}_{\mathrm{B}} = \sqrt{\beta_{\mathrm{B,L}}\beta_{\mathrm{E,L}}} e^{j\psi(\theta_{\mathrm{B}}, \theta_{\mathrm{E}})} \psi(\theta_{\mathrm{B}}, \theta_{\mathrm{E}}, M). \tag{9}$$

In the uplink training phase, legitimate user B transmits an uplink pilot symbol with the transmit power P_{B} and P_{E} is eavesdropper' power. At training time

slot j, the pilots sent by Bob and Eve are $\mathbf{p}_j{}^{\mathrm{x}} \in \mathbb{A}$ respectively, where $\mathbb{A}$ denotes the set of all training symbols. For fixed training scheme and most practical applications, the pilot set $\mathbb{A}$ used by Bob is publicly known and typically specified in the standard. Hence, in this case, Eve can transmit the same pilots as Bob. In this paper, the pilot alphabet A is assumed to be a PSK alphabet, with N-PSK symbols

$$\mathbb{A} = \{e^{im2\pi/N} : 0 \leq m \leq (N-1)\}. \quad (10)$$

The pre-processed received pilot signal at the base station is

$$\mathbf{y}_{\mathrm{j}} = \sqrt{\mathrm{P_B}}\mathbf{p}_{\mathrm{j}}^{\mathrm{B}}\mathbf{h}_{\mathrm{AB}} + \sqrt{\mathrm{P_E}}\mathbf{p}_{\mathrm{j}}^{\mathrm{E}}\mathbf{h}_{\mathrm{AE}} + \mathbf{n}_{\mathrm{j}}. \quad (11)$$

where $\mathbf{n}_j \sim \mathcal{CN}(\mathbf{0}, \sigma_j^2\mathbf{I}_{N_{\mathrm{t}}})$ is the additive white Gaussian noise and

$$\mathbf{h}_{\mathrm{AB}} = \mathbf{g}_{\mathrm{B}} + \beta_{\mathrm{B,N}}^{1/2}\mathbf{w}_{\mathrm{B}};\ \mathbf{h}_{\mathrm{AE}} = \mathbf{g}_{\mathrm{E}} + \beta_{\mathrm{E,N}}^{1/2}\mathbf{w}_{\mathrm{E}}. \quad (12)$$

Assume that the base station applies the linear minimum mean squared errors (MMSE) estimation method to obtain the following channel estimate. For the convenience, let $\mathrm{R_T^{11}}$ be a product defined on a $\mathbf{h}_{\mathrm{AB}}^{\mathrm{H}}\mathbf{h}_{\mathrm{AB}}$, $\mathrm{R_T^{12}}$ be a product defined on a $\mathbf{h}_{\mathrm{AB}}^{\mathrm{H}}\mathbf{h}_{\mathrm{AE}}$, $\mathrm{R_T^{21}}$ be a product defined on a $\mathbf{h}_{\mathrm{AE}}^{\mathrm{H}}\mathbf{h}_{\mathrm{AB}}$ and $\mathrm{R_T^{22}}$ be a product defined on a $\mathbf{h}_{\mathrm{AE}}^{\mathrm{H}}\mathbf{h}_{\mathrm{AE}}$.
After some manipulation, we obtain $\mathrm{R_T^{11}} = \beta_{\mathrm{B}}\mathrm{M}$; $\mathrm{R_T^{12}} = \sqrt{\beta_{\mathrm{B,N}}\beta_{\mathrm{E,N}}}\mathrm{e}^{\mathrm{j}\phi(\theta_{\mathrm{B}},\theta_{\mathrm{E}})}\alpha(\theta_{\mathrm{B}},\theta_{\mathrm{E}},\mathrm{M})$; $\mathrm{R_T^{21}} = [\mathrm{R_T^{12}}]^{\mathrm{H}}$; $\mathrm{R_T^{22}} = \beta_{\mathrm{E}}\mathrm{M}$.

III. Uplink Pilot Contamination Attack

Although it is difficult for BS to differentiate whether the pilots are from Bob only or contaminated by Eve, if BS has the knowledge of channel vector and they differ significantly, signal strength deviations from what is expected can be observed, and detection probability increases. Nevertheless, deterministic knowledge of Bob's pilots is detrimental for detection of Eve. Instead, if Bob transmits pilots randomly, then the probability of observing deviations from the expected signal increases. This observation forms the basis of our random pilot detection scheme, which is described next. We want to emphasize that our scheme does not need the knowledge of channel vector. The received signals during the two training slots are given by

$$\mathbf{y}_1 = \sqrt{P_{\mathrm{B}}}\mathbf{p}_1^{\mathrm{B}}\mathbf{h}_{\mathrm{A,B}} + \sqrt{P_{\mathrm{E}}}\mathbf{p}_1^{\mathrm{E}}\mathbf{h}_{\mathrm{A,E}} + \mathrm{n}_1, \quad (13)$$

$$\mathbf{y}_2 = \sqrt{P_{\mathrm{B}}}\mathbf{p}_2^{\mathrm{B}}\mathbf{h}_{\mathrm{A,B}} + \sqrt{P_{\mathrm{E}}}\mathbf{p}_1^{\mathrm{E}}\mathbf{h}_{\mathrm{A,E}} + \mathrm{n}_2. \quad (14)$$

Next, we form the product of the two received vectors

$$\mathbf{z}_{12} = \frac{1}{M}\mathbf{y}_1^{\mathrm{H}}\mathbf{y}_2. \quad (15)$$

We have:

$$\mathbf{z}_{12}^{\mathrm{E}} = \frac{1}{M}\Big\{P_{\mathrm{B}}(\mathbf{p}_1^{\mathrm{B}})^{\mathrm{H}}\mathbf{p}_2^{\mathrm{B}}\mathbf{R}_T^{11} + \sqrt{\mathrm{P_B P_E}}(\mathbf{p}_1^{\mathrm{B}})^{\mathrm{H}}\mathbf{p}_2^{\mathrm{E}}\mathbf{R}_{\mathrm{T}}^{12} + \sqrt{\mathrm{P_B P_E}}(\mathbf{p}_1^{\mathrm{E}})^{\mathrm{H}}\mathbf{p}_2^{\mathrm{B}}\mathbf{R}_{\mathrm{T}}^{12} + \mathrm{P_E}(\mathbf{p}_1^{\mathrm{E}})^{\mathrm{H}}\mathbf{p}_2^{\mathrm{E}}\mathbf{R}_{\mathrm{T}}^{22} + \mathbf{N}_{12}^{\mathrm{E}}\Big\}. \quad (16)$$

Where:

$$\mathbf{N}_{12}^{\mathrm{E}} = \sqrt{P_{\mathrm{B}}}(\mathbf{p}_1^{\mathrm{B}})^{\mathrm{H}}\mathbf{h}_{\mathrm{A,B}}^{\mathrm{H}}\mathbf{n}_2 + \sqrt{P_{\mathrm{E}}}(\mathrm{p}_1^{\mathrm{E}})^{\mathrm{H}}(\mathbf{h}_{\mathrm{A,E}})^{\mathrm{H}}\mathbf{n}_2 + \sqrt{P_{\mathrm{B}}}\mathbf{p}_2^{\mathrm{B}}\mathbf{h}_{\mathrm{A,B}}\mathbf{n}_1^{\mathrm{H}} + \sqrt{P_{\mathrm{E}}}\mathbf{p}_2^{\mathrm{E}}\mathbf{h}_{\mathrm{A,E}}\mathbf{n}_1 + \mathbf{n}_1^{\mathrm{H}}\mathbf{n}_2. \quad (17)$$

We offered a way to detect wiretapping device and calculated the probability of detecting E

A. Random pilot detection scheme: Noiseless case

To show this idea, we have first removed up to the noise in the received signal. If Eve is not active the cross product becomes.

$$\mathbf{z}_{12}^0 = \frac{1}{M}P_{\mathrm{B}}\left(\mathbf{p}_1^{\mathrm{B}}\right)^{\mathrm{H}}\mathbf{p}_2^{\mathrm{B}}\mathbf{h}_{\mathrm{A,B}}^{\mathrm{H}}\mathbf{h}_{\mathrm{A,B}}. \quad (18)$$

$$\mathbb{E}[\mathbf{z}_{12}^0] = \mathrm{P_B}\beta_{\mathrm{B}}(\mathrm{p}_1^{\mathrm{B}})^{\mathrm{H}}\mathrm{p}_2^{\mathrm{B}}. \quad (19)$$

- If E is not present, A receives z_{12} is a scaled N-PSK symbol.
- If E is present, A receives z_{12}^{E}. When z_{12}^{E} is a scaled N-PSK symbol,then E was not detected. This means the z_{12}^{E} lie one of the PSK lines.

Lemma 1. *For a couple $\mathbf{p}_j \in \mathbb{A}$ is N-PSK, then $\mathbf{p}_1^{\mathrm{H}}\mathbf{p}_2 \in N$-PSK*

Proof: At the training time slot j, pilots sent by Bob and Eve are are $\mathbf{p}_j{}^{\mathrm{B}} \in \mathbb{A}$ and $\mathbf{p}_j{}^{\mathrm{E}} \in \mathbb{A}$, $\mathbb{A}$ is N-PSK respectively, where $\mathbb{A}$ denotes the set of all training symbols. For fixed training scheme and most practical applications, the pilot set $\mathbb{A}$. The pilot alphabet A is assumed to be a PSK alphabet, with N -PSK symbols $\mathbb{A} = e^{in_j2\pi/N} : 0 \leq n_j \leq N-1$ $n_j \in (0, 1, ..., N-1)$. At time slot j = 1,2.
We have:

$$\mathbf{p}_1^{\mathrm{H}}\mathbf{p}_2 = A_1A_2e^{j(n_2-n_1)\frac{2\pi}{N}},$$
$$[n_2 - n_1] \in \{0, 1, ..., N-1\}$$
$$z = \arg\left|\mathbf{p}_1^{\mathrm{H}}\mathbf{p}_2\right| = (n_2 - n_1)\frac{2\pi}{N} \quad (20)$$

$\mathbf{z}$ is phase of a valid N-PSK ■
From Lemma 1 we have: $(\mathbf{p}_1^{\mathrm{B}})^{\mathrm{H}}\mathbf{p}_2^{\mathrm{B}}$ is phase of a valid N-PSK. Hence $\mathbb{E}[\mathbf{z}_{12}]$ to be a scaled PSK symbol. We have the discussion based on (15) and (19). If the Eve absents, then Alice receives $\mathrm{z_{12}^0} = \mathrm{y_1^H y_2}$,which is scaled PSK symbol. If Eve is present, then Alice receives $\mathrm{z_{12}^E} = \mathrm{y_1^H y_2}$ is not scaled PSK. So that Alice easy detects Eves present. For Eve to remain undetected, $\mathrm{z_{12}^E}$ have to be a scaled PSK symbol. That means $\mathrm{z_{12}^E}$ must lie in one of the $\mathrm{N}/2$ line.

With these observations, the detection procedure can be formulated as: if $\mathrm{y}_1^{\mathrm{H}}\mathrm{y}_2$ is on PSK line, Eve is absent; otherwise, Eve is present.
If Eve is active, the cross product becomes:

$$\begin{aligned}\mathbb{E}\left[\mathbf{z}_{12}^{\mathrm{E}}\right] =&\frac{1}{M}(\mathrm{p}_1^{\mathrm{B}})^{\mathrm{H}}\mathrm{p}_2^{\mathrm{B}}\Big\{P_{\mathrm{B}}\mathrm{M}\beta_{\mathrm{B}}\\ &+\sqrt{P_{\mathrm{B}}P_{\mathrm{E}}}\beta_{\mathrm{B,N}}^{1/2}\beta_{\mathrm{E,N}}^{1/2}\mathrm{e}^{\mathrm{j}\theta(\theta_{\mathrm{B}},\theta_{\mathrm{E}})}\alpha(\theta_{\mathrm{B}},\theta_{\mathrm{E}},\mathrm{N_t})\\ &\left(\mathbf{p}_2^{\mathrm{B}}\right)^{\mathrm{H}}\left(\mathbf{p}_2^{\mathrm{E}}\right)\\ &+\sqrt{P_{\mathrm{B}}P_{\mathrm{E}}}\beta_{\mathrm{B,N}}^{1/2}\beta_{\mathrm{E,N}}^{1/2}\mathrm{e}^{-\mathrm{j}\theta(\theta_{\mathrm{B}},\theta_{\mathrm{E}})}\alpha(\theta_{\mathrm{B}},\theta_{\mathrm{E}},\mathrm{N_t})\\ &\left(\mathbf{p}_1^{\mathrm{E}}\right)^{\mathrm{H}}\left(\mathbf{p}_1^{\mathrm{B}}\right)\\ &+P_{\mathrm{E}}\mathrm{M}\beta_{\mathrm{E}}\left(\mathbf{p}_2^{\mathrm{B}}\right)^{\mathrm{H}}\left(\mathbf{p}_1^{\mathrm{B}}\right)(\mathrm{p}_1^{\mathrm{E}})^{\mathrm{H}}\mathrm{p}_2^{\mathrm{E}}\Big\}.\end{aligned} \tag{21}$$

The product $(\mathrm{p}_1^{\mathrm{X}})^{\mathrm{H}}\mathrm{p}_2^{\mathrm{X}}$ is a scaled PSK symbol. So in the order for $\mathrm{z}_{12}^{\mathrm{E}}$ to be a scaled PSK symbol, the angle of the vector scalar product in (21) must equal the angle of some PSK symbol. If $\mathrm{p}_2^{\mathrm{E}} = \mathrm{p}_1^{\mathrm{E}}(\mathrm{p}_1^{\mathrm{B}})^{\mathrm{H}}\mathrm{p}_2^{\mathrm{B}}$. Else $\mathrm{p}_1^{\mathrm{E}}(\mathrm{p}_1^{\mathrm{B}})^{\mathrm{H}} \neq \mathrm{p}_2^{\mathrm{E}}(\mathrm{p}_2^{\mathrm{B}})^{\mathrm{H}}$, the angle of the above vector scalar product will, with probability one, not be equal to the angle of any PSK symbol. For the situation Eve to remain undetected, that mean Eve can guess the pilot $\mathrm{p}_1^{\mathrm{E}}(\mathrm{p}_1^{\mathrm{B}})^{\mathrm{H}}\mathrm{p}_2^{\mathrm{B}}$ at second time slot. Since Eve can guess the pilot $\mathrm{p}_1^{\mathrm{E}}(\mathrm{p}_1^{\mathrm{B}})^{\mathrm{H}}\mathrm{p}_2^{\mathrm{B}}$ is a random PSK symbol. But in the almost communication systems must take into account the noise. Then probability of detecting Eve will change and appears false alarm probability.

B. Random pilot detection scheme: Noisy case case

In the previous section we have assumed that noise was not present.Next we discuss the impact of noise in the system and detected probability of Eve.
Eve is not active in both slots. Then the received signals during the two training slots is

$$\begin{aligned}\mathbf{y}_1 =&\sqrt{P_{\mathrm{B}}}\mathbf{p}_1^{\mathrm{B}}\mathbf{h}_{\mathrm{A,B}} + \mathbf{n}_1\\ \mathbf{y}_2 =&\sqrt{P_{\mathrm{B}}}\mathbf{p}_2^{\mathrm{B}}\mathbf{h}_{\mathrm{A,B}} + \mathbf{n}_2.\end{aligned} \tag{22}$$

If Eve is not present in both time slot, so cross product y_{12}^0 becomes:

$$\mathrm{z}_{12}^0 = \frac{1}{\mathrm{M}}\left[\mathrm{P_B}(\mathrm{p}_1^{\mathrm{B}})^{\mathrm{H}}\mathrm{p}_2^{\mathrm{B}}\mathrm{h}_{\mathrm{A,B}}^{\mathrm{H}}\mathrm{h}_{\mathrm{A,B}} + \mathrm{N}_{12}^0\right]. \tag{23}$$

where

$$\begin{aligned}\mathrm{N}_{12}^0 = \frac{1}{\mathrm{M}}\Big[&\sqrt{\mathrm{P_B}}(\mathrm{p}_1^{\mathrm{B}})^{\mathrm{H}}\mathbf{h}_{\mathrm{A,B}}^{\mathrm{H}}\mathbf{n}_2 + \sqrt{\mathrm{P_B}}\mathrm{p}_2^{\mathrm{B}}\mathbf{n}_1^{\mathrm{H}}\mathbf{h}_{\mathrm{A,B}}\\ &+\mathbf{n}_1^{\mathrm{H}}\mathbf{n}_2\Big].\end{aligned} \tag{24}$$

is the equivalent noise. The mean of N_{12}^0 is $\mathbb{E}[\mathrm{N}_{12}^0] = 0$. we going to find the variance $\mathrm{S_E^0}$ of the interference variable N_{12}^0.

$$S_{\mathrm{E}}^0 = \frac{1}{\mathrm{M}}[2P_{\mathrm{B}}\mathrm{M}\beta_{\mathrm{B}}\sigma^2 + \sigma^4]. \tag{25}$$

In the absence of Eve, z_{12}^0 equals a scaled PSK symbol disturbed by complex Gaussian noise with zero mean and variance S_{E}^0.
On the other hand, if Eve is contaminating the pilots, we form the product of the two received vectors:

$$\begin{aligned}\mathbf{z}_{12}^{\mathrm{E}} = \frac{1}{M}\Big\{&\mathrm{P_B}\left(\mathbf{p}_1^{\mathrm{B}}\right)^{\mathrm{H}}\mathrm{p}_2^{\mathrm{B}}\mathbf{h}_{\mathrm{A,B}}^{\mathrm{H}}\mathbf{h}_{\mathrm{A,B}}\\ &+\sqrt{\mathrm{P_BP_E}}\left(\mathbf{p}_1^{\mathrm{B}}\right)^{\mathrm{H}}\left(\mathbf{p}_2^{\mathrm{E}}\right)(\mathbf{h}_{\mathrm{A,B}})^{\mathrm{H}}\mathbf{h}_{\mathrm{A,E}}\\ &+\sqrt{\mathrm{P_BP_E}}\left(\mathbf{p}_1^{\mathrm{E}}\right)^{\mathrm{H}}\left(\mathbf{p}_2^{\mathrm{B}}\right)\mathbf{h}_{\mathrm{A,E}}^{\mathrm{H}}\mathbf{h}_{\mathrm{A,B}}\\ &+\mathrm{P_E}\left(\mathbf{p}_1^{\mathrm{E}}\right)^{\mathrm{H}}\left(\mathbf{p}_2^{\mathrm{E}}\right)\mathbf{h}_{\mathrm{A,E}}^{\mathbf{H}}\mathbf{h}_{\mathrm{A,E}} + \mathrm{N}_{12}^{\mathrm{E}}\Big\}.\end{aligned} \tag{26}$$

where:

$$\begin{aligned}\mathbf{N}_{12}^{\mathrm{E}} =&\sqrt{P_{\mathrm{B}}}\mathbf{p}_1\mathbf{h}_{\mathrm{A,B}}\mathbf{n}_2 + \sqrt{P_{\mathrm{E}}}(\mathrm{p}_1^{\mathrm{B}})^{\mathrm{H}}(\mathbf{h}_{\mathrm{A,E}})^{\mathrm{H}}\mathbf{n}_2\\ &+\sqrt{P_{\mathrm{B}}}\mathbf{p}_2\mathbf{h}_{\mathrm{A,B}}\mathbf{n}_1^{\mathrm{H}} + \sqrt{P_{\mathrm{E}}}\mathbf{p}_2\mathbf{h}_{\mathrm{A,E}}\mathbf{n}_1 + \mathbf{n}_1^{\mathrm{H}}\mathbf{n}_2.\end{aligned} \tag{27}$$

$$\begin{aligned}\mathbb{E}[\mathbf{z}_{12}^{\mathrm{E}}] =&\mathbb{E}\Bigg[\frac{1}{M}\Big\{\mathrm{P_B}(\mathbf{p}_1^{\mathrm{B}})^{\mathrm{H}}\mathbf{p}_2^{\mathrm{B}}\mathbf{R}_T^{11}\\ &+\sqrt{\mathrm{P_BP_E}}(\mathbf{p}_1^{\mathrm{B}})^{\mathrm{H}}\mathbf{p}_2^{\mathrm{E}}\mathbf{R}_{\mathrm{T}}^{12}\\ &+\sqrt{\mathrm{P_BP_E}}(\mathbf{p}_1^{\mathrm{E}})^{\mathrm{H}}\mathbf{p}_2^{\mathrm{B}}\mathbf{R}_{\mathrm{T}}^{12} + \mathrm{P_E}(\mathbf{p}_1^{\mathrm{E}})^{\mathrm{H}}\mathbf{p}_2^{\mathrm{E}}\mathbf{R}_{\mathrm{T}}^{22}\Big\}\Bigg].\end{aligned} \tag{28}$$

For a given realization of the pilots and channels, the interference $\mathrm{N}_{12}^{\mathrm{E}}$ converges to a complex Gaussian variable with zero mean and variance $S_{\mathrm{E}}^{\mathrm{M}}$

$$\begin{aligned}S_{\mathrm{M}}^{\mathrm{E}} =&\sigma^2\Big\{2\mathrm{P_B}\mathrm{R}_{\mathrm{T}}^{11} + \sqrt{P_{\mathrm{B}}P_{\mathrm{E}}}[(\mathrm{p}_1^{\mathrm{B}})^{\mathrm{H}}\mathrm{p}_1^{\mathrm{E}} + (\mathrm{p}_2^{\mathrm{B}})^{\mathrm{H}}\mathrm{p}_2^{\mathrm{E}}]\mathrm{R}_{\mathrm{T}}^{12}\\ &+\sqrt{P_{\mathrm{B}}P_{\mathrm{E}}}[(\mathrm{p}_1^{\mathrm{E}})^{\mathrm{H}}\mathrm{p}_1^{\mathrm{B}} + (\mathrm{p}_2^{\mathrm{E}})^{\mathrm{H}}\mathrm{p}_2^{\mathrm{B}}]\mathrm{R}_{\mathrm{T}}^{21}\\ &+2P_{\mathrm{E}}\mathrm{R}_{\mathrm{T}}^{22} + \sigma^2\Big\}.\end{aligned} \tag{29}$$

Applying the same analysis in previous section, If $\mathrm{p}_1^{\mathrm{E}}(\mathrm{p}_1^{\mathrm{B}})^{\mathrm{H}} = \mathrm{p}_2^{\mathrm{E}}(\mathrm{p}_2^{\mathrm{B}})^{\mathrm{H}}$, then $\mathrm{z}_{12}^{\mathrm{E}}$ will be equal to a scaled PSK symbol plus $\mathrm{N}_{12}^{\mathrm{E}}$. Hence, in this case, the situation is similar as that in the absence of Eve, and the probability of detecting Eve decreases. On the other hand, if $\mathrm{p}_1^{\mathrm{E}}(\mathrm{p}_1^{\mathrm{B}})^{\mathrm{H}} \neq \mathrm{p}_2^{\mathrm{E}}(\mathrm{p}_2^{\mathrm{B}})^{\mathrm{H}}$, $\mathrm{z}_{12}^{\mathrm{E}}$ will be equal to a symbol difference from a PSK symbol plus N_{12}^{E}. The variance of $\mathrm{N}_{12}^{\mathrm{E}}$ will vary in the same way as in the previous case.

C. Construction of Detection Regions

The results in the previous sections show us how to construct the detection regions, i.e the regions in which BS decides whether Eve is contamination or not, depending on if the scalar product $\mathrm{y}_1^{\mathrm{H}}\mathrm{y}_2/\mathrm{M}$ is outside or inside the detection region, respectively. Since the scalar product z_{12} in (23) equal the sum of a PSK symbol scalar scaled with $C_{\mathrm{B}} = P_{\mathrm{B}}\beta_{\mathrm{B}}$ and Gaussian noise, BS decides that Eve is not contaminating if the

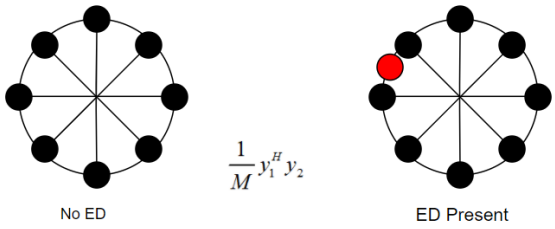

Fig. 2. Detection scheme of the random pilot detection scenario. First B transmits random PSK symbols. After the processing at BS, if E absent so the product of the two received signals should be a valid PSK symbol. Otherwise E present.

scalar product $\mathrm{y}_1^{\mathrm{H}}\mathrm{y}_2/\mathrm{M}$ is within a certain distance $r(C_{\mathrm{B}})$ from some PSK line. $r(C_{\mathrm{B}})$ needs to increase with the scaling $r(C_{\mathrm{B}})$, because the variance $\mathrm{S}_{\mathrm{M}}^0$ of the Gaussian noise N_{12}^0 increases with C_{B}. From the signal space perspective, Gaussian noise corresponds to a circle centered around 0 with radius $\sqrt{S_{\mathrm{M}}^0}$. This property leads us to construct $r(C_{\mathrm{B}})$ In the order to detect Eve, Alice performs the following procedure for some realizations of the scalar product

$$s_M^0 = \frac{N_0}{N_t^2}(MN_t + 2P_B\mathbf{R_T^{11}}) \tag{30}$$

To detect Eve, BS follows the algorithm below:
With each corresponding real part of $y_{12} = \frac{(y_1^H y_2)}{N_t}$
Calculate $|y_{12} - xv|$ with $r(x)$ Situations happened:
Case 1: $|y_{12} - xv| > r(x)$ Alice announces that Eve is absent
Case 2: $|y_{12} - xv| < r(x)$ Alice announces that Eve had presented

IV. Numerical Results

To evaluate the performance of our detection scheme, we simulate the detection probability and the false-alarm probability. False-alarm probability is defined as the probability of detecting Eve, given that Eve is not present. We considered a network has only one cell, in which the BS is located at center in the cell and legitimated user Bob (B) and eavesdropper device (Eve) arranged randomly in the cell. Assuming the effect of shadowing is ignored, then large-scale fading follows the [16]

$$\beta_{\mathrm{X,Y}} = 32.4 + 10n_{\mathrm{Y}}\log_{10}(d_{\mathrm{3D,X}}) + 20\log_{10}(f_{\mathrm{c}}).$$

where $X \in \mathcal{X}$, $Y \in \mathcal{Y} = \{\mathrm{L}, \mathrm{N}\}$, $d_{\mathrm{3D,X}}$ is the distance in meters from base station to node X in 3-D space,$f_{\mathrm{c}} = 3.5$GHz is the carrier frequency,n_{Y} is the exponential coefficient of transmission (PLE: path-loss exponent). Moreover $d_{\mathrm{3D,X}}$ follows the formula $d_{\mathrm{3D,X}} = \sqrt{d_{\mathrm{2D,X}}^2 + (h_{\mathrm{A}} - h_{\mathrm{X}})^2}$, where $d_{\mathrm{2D,X}}$ is the distance from the base station to the node X in the 2-D space, h_{A} is height of the base station A, and h_{X} is height of the node X [16]. For general, suppose $h_{\mathrm{A}} =$ 10m and $h_{\mathrm{B}} = h_{\mathrm{E}} = 1.5$m. The paper investigates the urban cell environment (UMa: Urban Macro), then $n_{\mathrm{L}} = 2$ for LOS and $n_{\mathrm{N}} = 2.9$ for NLOS. Follow [16], for UMa environment then κ measured in dB is a Gaussian random variable $\mathcal{N}(9, 3.5)$. For simplicity, we assume $\kappa_{\mathrm{B}} = \kappa_{\mathrm{E}} = 9$dB. Assuming the system works at bandwidth 10MHz,the transmit power of the base station is $p_{\mathrm{d}} = 2$dB. Assuming the distance between adjacent antennas at the base station is half wavelength,that is $d = 0.5$. First, we consider a simulation scenario , while Eve or node E closes to the legitimated user or node B. Some simulated parameters of this scenario are: i) distance from node E and node B to base station 300m, ii) channel model coefficients Rician is $\kappa_{\mathrm{B}} = \kappa_{\mathrm{E}} = 9$dB, and iii) Simulation results are averaged over 150000 samples.

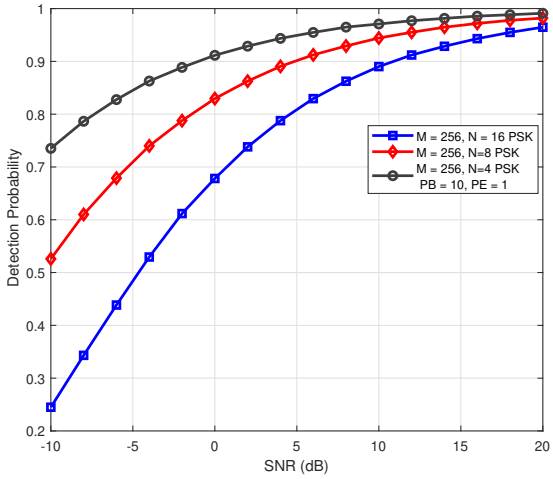

Fig. 3. Detection probabilities vs. SNR for $\mathrm{M} = 256$, $\mathrm{P_B} = 10$, $\mathrm{P_E} = 1$, $\Phi_{\mathrm{B}} = 0.1$ rad and $\Phi_{\mathrm{E}} = 0.1$ rad

Figure 3 shows the detection probability v.s SNR of this scenario for 256 antennas at BS, $\mathrm{P_B} = 10$ dB $\mathrm{P_E} = 1$ dB, $\mathrm{N} = [4, 8, 16]$- PSK. There, SNR is defined as $\mathrm{SNR} = \frac{\mathrm{P_B}}{\mathrm{N_0}}$ in dB. As expected, the detection probability increases with SNR; in the high SNR domain, detection probability go to 1. In [7] pointed out that Eve's transmit power is larger than legitimated user Bob's then Eve easier to detect. But in Massive MIMO channels with Rician fading and larger number of antennas not difficult to detect Eve. Figure 4 shown detection probability v.s SNR of our system for N = 8 PSK and different number of antennas at BS. As expected, the detection probability increases with SNR and go to 1 when SNR increase. Figure 5 shows the false-alarm probabilities decreases even while $\Phi_{\mathrm{E}} = 0.1$ rad, $\Phi_{\mathrm{B}} = 0.1$ rad with SNR = [-1, 1, 5, 10] dB when number of antennas increases.

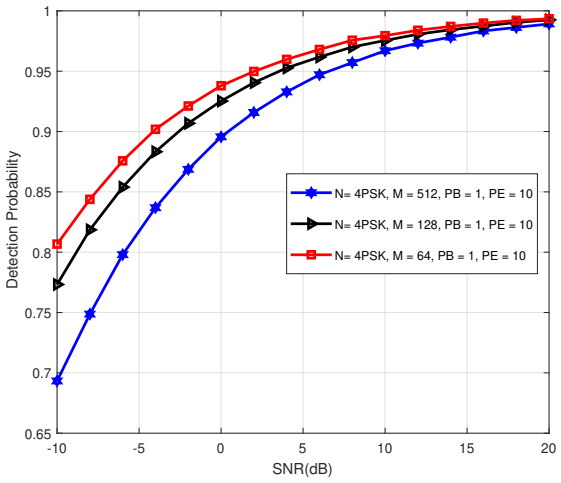

Fig. 4. Detection probabilities vs. SNR for N = 8-PSK, $P_B = 10$, $P_E = 1$ and $\Phi_B = 0.1$ rad và $\Phi_E = 0.1$ rad, d_B= 300 m, d_E = 200 m

.

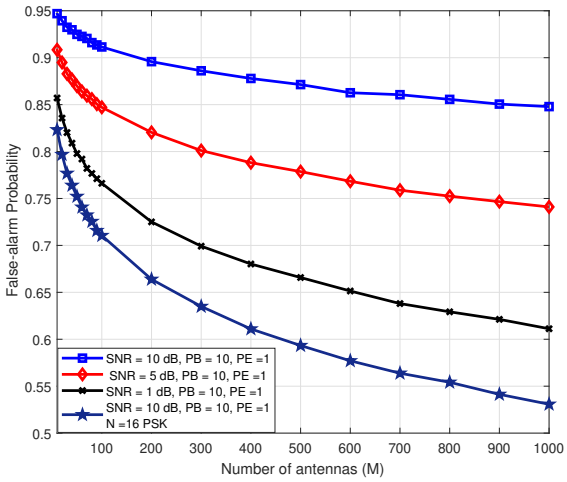

Fig. 5. False-alarm probabilities $\Phi_E = 0.1$ rad, $\Phi_B = 0.1$ rad, SNR = 3 dB vs. M

.

V. Conclusion and Future Work

In this paper, we presented detection schemes based on random PSK pilots. The detection schemes require only two training slots to perform detection at the base station without any prior channel knowledge. Numerical results have shown that the detection scheme provides a high detection probability. Future work we investigate affect of passive eavesdropper in Massive MIMO who randomly present on any training slot in spatially-ucorrelated and spatially-uncorrelated Rician channels.

References

[1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. on Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, 2010.

[2] J. Zhang, J. Fan, B. Ai, and D. W. K. Ng, "Noma-based cell-free massive mimo over spatially correlated rician fading channels," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.

[3] L. Sanguinetti, E. Björnson, and J. Hoydis, "Toward massive mimo 2.0: Understanding spatial correlation, interference suppression, and pilot contamination," *IEEE Transactions on Communications*, vol. 68, no. 1, pp. 232–257, 2020.

[4] I. F. Akyildiz, A. Kak, and S. Nie, "6g and beyond: The future of wireless communications systems," *IEEE Access*, vol. 8, pp. 133 995–134 030, 2020.

[5] D. Kapetanovic, G. Zheng, and F. Rusek, "Physical layer security for massive MIMO: An overview on passive eavesdropping and active attacks," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 21–27, 2015.

[6] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proc. of the IEEE*, vol. 104, no. 9, pp. 1727–1765, Sep. 2016.

[7] D. Kapetanović, G. Zheng, K. Wong, and B. Ottersten, "Detection of pilot contamination attack using random training and massive MIMO," in *Proc. of IEEE Int. Symp. Personal, Indoor, Mobile Radio Commun. (PIMRC)*, Sep. 2013, pp. 13–18.

[8] T. Yang, R. Zhang, X. Cheng, and L. Yang, "Secure massive MIMO under imperfect CSI: Performance analysis and channel prediction," *to appear in IEEE Trans. Info. Forensics Security*, 2018.

[9] X. Zhang, D. Guo, and K. Guo, "Secure performance analysis for multi-pair AF relaying massive MIMO systems in Ricean channels," *IEEE Access*, vol. 6, pp. 57 708–57 720, 2018.

[10] A. Mukherjee and A. Swindlehurst, "A full-duplex active eavesdropper in MIMO wiretap channels: Construction and countermeasures," in *Proc. IEEE Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, U.S.A., Nov. 2011, pp. 265–269.

[11] D. B. Rawat, K. Neupane, and M. Song, "A novel algorithm for secrecy rate analysis in massive MIMO system with target SINR requirements," in *Proc. of IEEE Int. Conf. Computer Commun. (INFOCOM)*, Apr. 2016, pp. 53–58.

[12] O. Ozdogan, E. Bjornson, and E. G. Larsson, "Massive MIMO with spatially correlated rician fading channels," *Submitted to IEEE Trans. Commun.*, 2018.

[13] V. L. Q. Giang and T. T. Kien, "Secret capacity of massive mimo with a passive eavesdropper," *Journal of Research and Development on Information and Communication Technology*, vol. V-3, no. 40, p. 1, 12 2018.

[14] C.-Y. Yeh and E. W. Knightly, "Feasibility of passive eavesdropping in massive MIMO: An experimental approach," in *Proc. of IEEE Conf. Commun. Network Security (CNS)*, Beijing, China, May 2018.

[15] A. Forenza, D. J. Love, and R. W. Heath, "Simplified spatial correlation models for clustered MIMO channels with different array configurations," *IEEE Trans. Veh. Tech.*, vol. 56, no. 4, pp. 1924–1934, Jul 2007.

[16] 3GPP TR 38.901, "Study on channel model for frequencies from 0.5 to 100 GHz," 3GPP, Technical Report v.15.0.0, Jun. 2018.