Detection of Pilot Contamination Attack Using Random Training and Massive MIMO

Dževdan Kapetanović*, Gan Zheng*, Kai-Kit Wong† and Björn Ottersten*
*Interdisciplinary Center for Security, Reliability and Trust (SnT)
University of Luxembourg, Luxembourg. Email: {dzevdan.kapetanovic, gan.zheng, bjorn.ottersten}@uni.lu.
†Department of Electronic and Electrical Engineering
University College London, UK. Email: kai-kit.wong@ucl.ac.uk.

***Abstract*—Channel estimation attacks can degrade the performance of the legitimate system and facilitate eavesdropping. It is known that pilot contamination can alter the legitimate transmit precoder design and strengthen the quality of the received signal at the eavesdropper, without being detected. In this paper, we devise a technique which employs random pilots chosen from a known set of phase-shift keying (PSK) symbols to detect pilot contamination. The scheme only requires two training periods without any prior channel knowledge. Our analysis demonstrates that using the proposed technique in a massive MIMO system, the detection probability of pilot contamination attacks can be made arbitrarily close to 1. Simulation results reveal that the proposed technique can significantly increase the detection probability and is robust to noise power as well as the eavesdropper's power.**

I. Introduction

There has been growing interest in physical layer security since the introduction of the degraded wiretap channel and secrecy capacity by the seminal work of Wyner [1]. Since then, much work has been reported in the literature from the viewpoints of both information theory [2]–[4] and signal processing [5]–[10]. Deviating from the main research efforts on secrecy capacity without considering on how channel state information (CSI) is obtained, we are interested the practical problem of detecting the attack on the channel estimation process from an eavesdropper. More specifically, we consider a multiple-antenna (MIMO) precoding system, where the eavesdropper (Eve) wants to overhear the communication from the legitimate transmitter (Alice) to the intended receiver (Bob), but does not attack Bob directly by, for instance, jamming. Thus, detecting Eve becomes an important yet difficult task.

The impact of CSI on secrecy has been recently investigated. A hybrid half-duplex adversary who either jams or eavesdrops at a given time based on different level of CSI was studied in [11], which illustrates that the main CSI is more valuable for the adversary than the jamming CSI in both delay-limited and ergodic scenarios. Intriguing counter-intuitive results were given in [12] showing that more knowledge to the eavesdroppers makes them more conservative in their attacks, and thus less harmful and similarly, providing more knowledge to the legitimate transmitter makes it more careful and less willing to transmit, which reduces the expected secrecy capacity.

The need of channel estimation for MIMO precoding systems has put them at risk of active attack on channel estimation. One such active eavesdropping attack was discussed in [13], where Eve contaminates the channel estimation phase. In particular, with a time-division-duplex (TDD) system, where downlink and uplink channels are considered reciprocal, Eve can inject the same pilot as Bob in the uplink channel training phase. If Eve has sufficient transmit power, it can control the channel estimation result at Alice; this will then change the matched filter precoding used in the downlink transmission phase. As such, Eve will not only strengthen its receive signal but also degrade the signal quality at Bob, which is extremely harmful for precoding systems. Worse, Eve's attack cannot be detected since its behavior is exactly the same as Bob's.

In this paper, we propose a scheme to detect the presence of Eve who attacks on the channel estimation using the method in [13]. The main idea is to utilize random pilots for channel estimation. We do not use pilot symbols as secret keys, which are normally known in the standard, but rather assume the set of pilot symbols to be publicly known. More specifically, we use phase-shift keying (PSK) symbols as the pilot symbols which are transmitted randomly. The scalar product between the received vectors is exploited to detect the presence of Eve.

First, we describe our scheme for any number of antennas at the base station, and derive a geometric region in the complex plane based on Gaussian approximation that can be used for detection. Then the potential of the emerging very large-scale MIMO (a.k.a. massive MIMO technique) is investigated for our scheme. Massive MIMO has received great attention for its impressive gain in throughput and energy efficiency [14]. Interested readers are referred to a tutorial in [15]. One property particularly useful to our detection algorithm is that with massive MIMO, the received thermal noise at Alice in the uplink can be averaged out so that the detection can be much simplified with improved performance. We illustrate that with massive MIMO and large constellation size, Eve can be detected with a probability arbitrarily close to one.[1]

The rest of this paper is organized as follows. In Section II, we introduce our system model. Section III presents our detection algorithm based on random training along with a theoretical analysis of its behavior in the presence of received noise. Based on the theoretical analysis, Section IV constructs

[1]It is worth noting that when Eve is completely passive without transmitting any signal, detecting its existence is more difficult. A method was proposed in [16] to detect passive eavesdropping from the local oscillator power leaked from Eve's RF front end. This is beyond the scope of this paper.

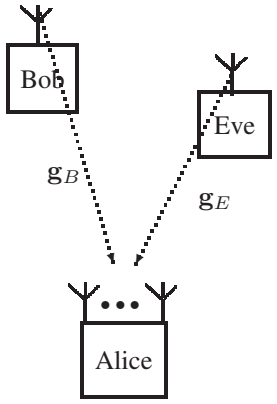

Fig. 1. Alice is a multi-antenna base station sending messages to the intended user Bob but Eve is the malicious single-antenna eavesdropper.

the detection regions. Section V studies the impact of massive MIMO on the detection. In Section VI, simulation results are provided and Section VII concludes the paper.

II. System Model

Consider a TDD system with a *base station* Alice, equipped with M antennas and (possibly) several single-antenna users, as depicted in Figure 1. Due to TDD, it is sufficient to focus on one *intended user* Bob, awaiting information from Alice. Alice adapts its precoder design to match the channel to Bob for enhanced performance. Due to reciprocity, channel estimation can be done in the uplink where Bob sends pilot symbols to Alice. There exists a malicious single-antenna terminal Eve to overhear the communication between Alice and Bob.

Let $\mathbf{g}_B = d_B\mathbf{h}_B$ be the $M \times 1$ vector channel from Bob to Alice and $\mathbf{g}_E = d_E\mathbf{h}_E$ be the $M \times 1$ vector channel from Eve to Alice. The elements in $\mathbf{h}_B$ and $\mathbf{h}_E$ are independent complex Gaussians with zero mean and unit variance, i.e., they represent small-scale fading. The scalars d_B and d_E are the large-scale fading coefficients for path loss and shadowing.

Understanding the importance of uplink channel estimation, instead of passively listening to the legitimate transmission, a more effective strategy for Eve is to contaminate the channel estimation phase. Next we briefly review a pilot contamination attack proposed in [13]. At training time slot j, the pilots sent by Bob and Eve are $p_j^B \in \mathcal{A}$ and $p_j^E \in \mathcal{A}$, respectively, where $\mathcal{A}$ denotes the set of all training symbols. For fixed training scheme and most practical applications, the pilot set $\mathcal{A}$ used by Bob is publicly known and typically specified in the standard. Hence, in this case, Eve can transmit the same pilots as Bob, i.e., $p_j^B = p_j^E = p_j$. The received signal at Alice becomes

$$\begin{aligned}\mathbf{y}_j &= \sqrt{P_B}p_j d_B\mathbf{h}_B + \sqrt{P_E}p_j d_E\mathbf{h}_E + \mathbf{n}_j \\ &= (\sqrt{P_B}d_B\mathbf{h}_B + \sqrt{P_E}d_E\mathbf{h}_E)p_j + \mathbf{n}_j, \end{aligned} \tag{1}$$

where each element in $\mathbf{n}_j$ is a complex Gaussian random variable with zero mean and variance N_0. Furthermore, P_B and P_E denote the average training power used by Bob and Eve, respectively. We assume that the large-scale fading coefficients d_B and d_E are unknown in advance to Alice. From (1), it is impossible for Alice to decide whether the received signal strength is due to Bob's channel only or Eve's as well. If Eve's channel is strong or its training power is high, it can dominate the channel estimation at Alice and even worse, during the transmission phase, Alice will use a precoder that adapts to this erroneous channel estimate. Thus, it is necessary for Alice to detect the presence of Eve and pause any transmission to Bob during Eve's presence.[2]

The work in [13] only analyzed error rate performance for the pilot contamination attack, and outlined a detection scheme (without analysis) that measures the variance of the received signal at Alice during a sufficiently long training period. Instead, we propose an efficient detection scheme that achieves plausible performance under the pilot contamination attack with only two training slots. We also rigorously characterize the performance of the proposed detection algorithm.

III. Random Pilot Detection Scheme

Although it is difficult for Alice to differentiate whether the pilots are from Bob only or contaminated by Eve, if Alice has the knowledge of d_B and d_E, 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 d_B and d_E.

A. Random Pilot Detection Scheme: Noiseless Case

To illustrate the idea, let us first discard the noise in the received signal. The pilot alphabet $\mathcal{A}$ is assumed to be a PSK alphabet, with N PSK symbols $\mathcal{A} = \{e^{i2\pi k/N} : 0 \le k \le N-1\}$. We assume that 2 training slots are used. Suppose that Eve is active in both slots. Then the received signals during the two training slots are, respectively, given by

$$\begin{cases} \mathbf{y}_1 = \sqrt{P_B}p_1^B d_B\mathbf{h}_B + \sqrt{P_E}p_1^E d_E\mathbf{h}_E, \\ \mathbf{y}_2 = \sqrt{P_B}p_2^B d_B\mathbf{h}_B + \sqrt{P_E}p_2^E d_E\mathbf{h}_E. \end{cases} \tag{2}$$

Next, we form the (scaled) inner product of the two received vectors:

$$\begin{aligned} z_{12}^E \triangleq \frac{\mathbf{y}_1^*\mathbf{y}_2}{M} = \frac{1}{M}&\left(\sqrt{P_B}p_1^B d_B\mathbf{h}_B + \sqrt{P_E}p_1^E d_E\mathbf{h}_E\right)^* \\ &\left(\sqrt{P_B}p_2^B d_B\mathbf{h}_B + \sqrt{P_E}p_2^E d_E\mathbf{h}_E\right) \end{aligned} \tag{3}$$

where the superscript $(\cdot)^*$ denotes Hermitian conjugate. Instead, if Eve is not active, the cross product $\mathbf{y}_1^*\mathbf{y}_2/M$ becomes

$$z_{12} \triangleq P_B\frac{(p_1^B)^*(p_2^B)d_B^2\|\mathbf{h}_B\|^2}{M}. \tag{4}$$

[2]Note that this argument assumes perfect synchronization of Eve and Bobs' transmissions. Any synchronization imperfections could potentially be useful in the detection process, but they are out of scope in this work.

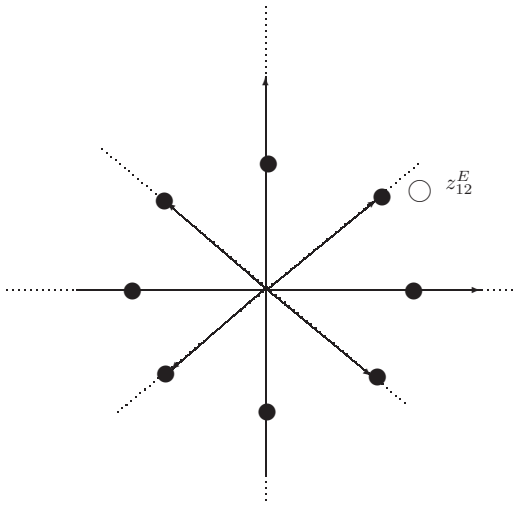

Fig. 2. Geometrical interpretation of the random pilot detection scheme. In the figure, Eve is detected since z_{12}^E is not located on one of the 4 dotted lines defined by a PSK symbol and the origin.

We have the following discussion based on (3) and (4).

- In the absence of Eve, Alice receives $z_{12} = \mathbf{y}_1^* \mathbf{y}_2 / M$, which is a scaled PSK symbol.
- If Eve is present, then Alice receives $z_{12}^E = \mathbf{y}_1^* \mathbf{y}_2 / M$. Hence, for Eve to remain undetected, z_{12}^E must be a scaled PSK symbol. Geometrically, this means that z_{12}^E must lie in one of the $N/2$ lines (henceforth called *PSK lines*) defined by the origin and each PSK symbol. See Figure 2.

With these observations in mind, the detection procedure can be formulated as: if $\mathbf{y}_1^* \mathbf{y}_2$ is on a PSK line, Eve is absent; otherwise, Eve is present. The following theorem characterizes the detection performance of the proposed scheme.

Theorem 1. *In the absence of noise, Eve can be detected with probability* $1 - \frac{1}{N}$.

Proof: It follows that

$$z_{12}^E = \frac{1}{M}(p_1^B)^* p_2^B (\sqrt{P_B} d_B \mathbf{h}_B + \sqrt{P_E} p_1^E (p_1^B)^* d_E \mathbf{h}_E)^* (\sqrt{P_B} d_B \mathbf{h}_B + \sqrt{P_E} p_2^E (p_2^B)^* d_E \mathbf{h}_E). \quad (5)$$

The product $\sqrt{P_B}(p_1^B)^* p_2^B$ is a scaled PSK symbol. Hence, in order for z_{12}^E to be a scaled PSK symbol, the angle of the vector scalar product in (5) must equal the angle of some PSK symbol. Due to the randomness of $\mathbf{h}_B$ and $\mathbf{h}_E$, if $p_1^E (p_1^B)^* \neq p_2^E (p_2^B)^*$, the angle of the above vector scalar product will, with probability one, not be equal to the angle of any PSK symbol. As such, z_{12}^E will not be a scaled PSK symbol with probability one. Instead, if $p_1^E (p_1^B)^* = p_2^E (p_2^B)^*$, z_{12}^E will be a scaled PSK symbol for any realization of $\mathbf{h}_B$ and $\mathbf{h}_E$. Hence, for Eve to remain undetected, Eve's pilot in the second time slot must equal $p_2^E = p_1^E (p_1^B)^* p_2^B$. Since $p_1^E (p_1^B)^* p_2^B$ is a random PSK symbol, Eve must guess the pilot $p_1^E (p_1^B)^* p_2^B$. Thus, the probability of Eve remaining undetected is $1/N$, i.e., the detection probability is $1 - 1/N$. ∎

Theorem 1 shows that by increasing the alphabet size, the detection probability converges to 1 if no noise is present. Note that no information about the large-scale fading coefficients is assumed, and this holds throughout this paper.

B. Random Pilot Detection Scheme: Noisy Case

So far, we have assumed that noise was not present. Next, the impact of noise, hence the model in (1), is considered.

If Eve is not contaminating, the scalar product z_{12} becomes

$$z_{12} = \frac{P_B d_B^2 \|\mathbf{h}_B\|^2}{M} (p_1^B)^* p_2^B + n_{12}. \quad (6)$$

where

$$n_{12} = \sqrt{P_B} d_B (p_1^B)^* \frac{\mathbf{h}_B^* \mathbf{n}_2}{M} + \sqrt{P_B} d_B p_2^B \frac{\mathbf{n}_1^* \mathbf{h}_B}{M} + \frac{\mathbf{n}_1^* \mathbf{n}_2}{M} \quad (7)$$

is the equivalent noise. The distribution of n_{12} is complicated and in the following lemma we study its property in the large antenna case.

Lemma 1. *For a given realization of* $\mathbf{h}_B$*, define*

$$s_M \triangleq \frac{N_0}{M^2}(MN_0 + 2P_B d_B^2 \|\mathbf{h}_B\|^2).$$

Then

$$\lim_{M \to \infty} \frac{n_{12}}{\sqrt{s_M}} \xrightarrow{d} \mathcal{CN}(0,1).$$

Proof: Due to space limitation, we sketch a brief outline of the proof only. First of all, write

$$n_{12} = \frac{1}{M}(\mathbf{n}_1 + \sqrt{P_B} d_B (p_1^B) \mathbf{h}_B)^* (\mathbf{n}_2 + \sqrt{P_B} d_B p_2^B \mathbf{h}_B) - P_B d_B^2 (p_1^B)^* p_2^B \|\mathbf{h}_B\|^2. \quad (8)$$

The scalar product in (8) is between two independent Gaussian vectors, with means $\sqrt{P_B} d_B (p_1^B) \mathbf{h}_B$ and $\sqrt{P_B} d_B p_2^B \mathbf{h}_B$, and both with variance N_0, respectively. Hence, the mean of n_{12} is $\mathbb{E}\{n_{12}\} = 0$. Also, n_{12} is a sum of M complex-valued normal product Gaussian variables. It can be shown that the Lyapunov condition in the Lyapunov central limit theorem is satisfied for the forth moment for this sum. Thus, we can conclude that in the limit $M \to \infty$, this sum converges to a complex Gaussian random variable with mean 0 and variance s_M. ∎

Lemma 1 shows that for a fixed channel realization, the interference variable n_{12} converges (in distribution) to a complex Gaussian variable with zero mean and variance s_M. Results in Figure 3 verify that this approximation is justified for as few as $M = 5$ antennas. Combining (6) and Lemma 1, it holds that in the absence of Eve, z_{12} equals a scaled PSK symbol disturbed by complex Gaussian noise with zero mean and variance s_M.

On the other hand, if Eve is contaminating the pilots, the cross product z_{12}^E equals

$$z_{12}^E = \frac{1}{M}\left(\sqrt{P_B} p_1^B d_B \mathbf{h}_B + \sqrt{P_E} p_1^E d_E \mathbf{h}_E\right)^* \left(\sqrt{P_B} p_2^B d_B \mathbf{h}_B + \sqrt{P_E} p_2^E d_E \mathbf{h}_E\right) + n_{12}^E, \quad (9)$$

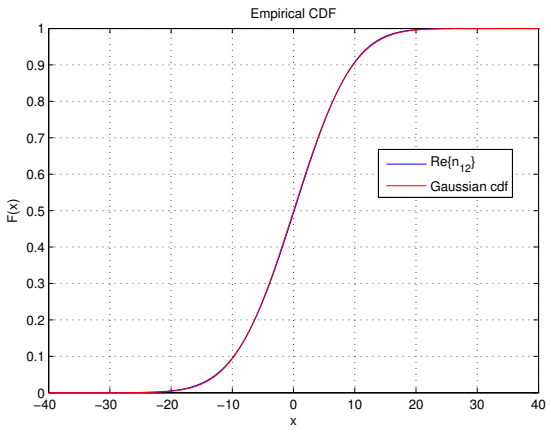

Fig. 3. The cumulative density functions (cdf) of the real-valued part $\mathrm{Re}\{n_{12}\}$ of n_{12} and a Gaussian cdf with mean 0 and variance s_M in Lemma 1. In this plot, $P_B d_B^2 \|\mathbf{h}_B\|^2 = 713.1196$ and $M = 5$. The complex-valued part of n_{12} gives the same cdf as the real-valued part. As seen, the empirical cdf of $\mathrm{Re}\{n_{12}\}$ coincides with a Gaussian cdf, as predicted by Lemma 1.

where the equivalent noise is

$$n_{12}^E = \frac{1}{M}\Big(\sqrt{P_B}d_B(p_1^B)^*\mathbf{h}_B^*\mathbf{n}_2 + \sqrt{P_E}d_E(p_1^E)^*\mathbf{h}_E^*\mathbf{n}_2 + \sqrt{P_B}d_B p_2^B\mathbf{n}_1^*\mathbf{h}_B + \sqrt{P_E}d_E p_2^E\mathbf{n}_1^*\mathbf{h}_E + \mathbf{n}_1^*\mathbf{n}_2\Big). \quad (10)$$

Similar to Lemma 1, we can prove the following result.

Lemma 2. *Define*

$$s_M^E \triangleq \frac{N_0}{M^2}(MN_0 + \|\sqrt{P_B}d_B p_1^B\mathbf{h}_B + \sqrt{P_E}d_E p_1^E\mathbf{h}_E\|^2 + \|\sqrt{P_B}d_B p_2^B\mathbf{h}_B + \sqrt{P_E}d_E p_2^E\mathbf{h}_E\|^2). \quad (11)$$

For a fixed realization of the pilots and the channels,

$$\lim_{M\to\infty} \frac{n_{12}^E}{\sqrt{s_M^E}} \xrightarrow{d} \mathcal{CN}(0,1). \quad (12)$$

Lemma 2 shows that for a given realization of the pilots and the channels, the interference n_{12}^E converges (in distribution) to a complex Gaussian variable with zero mean and variance s_M^E. As with n_{12}, n_{12}^E can be approximated very well with its limiting distribution for as few as 5 antennas.

Applying the same analysis in the proof of Theorem 1, if $p_1^E(p_1^B)^* = p_2^E(p_2^B)^*$, then z_{12}^E will be equal to a scaled PSK symbol plus n_{12}^E. The variance of n_{12}^E can be larger or smaller than the variance of n_{12}, depending on the pilots and the channel realizations. 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 $p_1^E(p_1^B)^* \neq p_2^E(p_2^B)^*$, z_{12}^E will be equal to a symbol different from a PSK symbol plus n_{12}^E; the variance of n_{12}^E will vary in the same way as in the previous case. However, the probability of detecting Eve will now increase, since the probability of z_{12}^E not being a scaled PSK symbol is larger.

IV. Construction of Detection Regions

The results in the previous section suggest how to construct the *detection regions*, i.e., the regions in which Alice decides whether Eve is contaminating or not, depending on if the scalar product $\mathbf{y}_1^*\mathbf{y}_2/M$ is outside or inside the detection region, respectively. Since the scalar product z_{12} in (6) equals the sum of a PSK symbol scaled with $c_B = P_B d_B^2\|\mathbf{h}_B\|^2/M$ and Gaussian noise, Alice decides that Eve is not contaminating if the scalar product $\mathbf{y}_1^*\mathbf{y}_2/M$ is within a certain distance $r(c_B)$ from some PSK line. $r(c_B)$ needs to increase with the scaling c_B, because the variance s_M of the Gaussian noise n_{12} increases with c_B. From the signal space perspective, Gaussian noise corresponds to a circle centered around 0 with radius $\sqrt{s_M}$. This property leads us to construct $r(c_B)$ as[3]

$$r(c_B) = c\frac{\sqrt{N_0(MN_0 + 2c_B)}}{M}, \quad (13)$$

Different values of the constant c will give smaller or larger detection regions. This parameter is fine-tuned in Section VI in order to achieve the best performance.

In order to detect Eve, Alice performs the following procedure for each realization of the scalar product $y_{12} = \mathbf{y}_1^*\mathbf{y}_2/M$:

Set $p = 1$.
For each PSK symbol $q \in \mathcal{A}$
  Define $f(x) \triangleq |y_{12} - xq| - r(x)$.

  If there is an x such that $f(x) < 0$, set $p = 0$.
  Otherwise continue.
End

If $p = 1$ after the above procedure completes, Alice decides that Eve is present; otherwise, it decides that Eve is not present. The geometrical interpretation of this procedure is to check whether there is a PSK symbol scaled with x, for which y_{12} is inside a circle with radius $r(x)$ centered at this scaled PSK symbol. In other words, it checks whether y_{12} is inside the detection region defined by the distance function $r(x)$.

V. Enhanced Detection Using Massive MIMO

Massive MIMO has received increasing attention lately. The main idea is to let the number of antennas at Alice, M, go to infinity. This is especially useful if unwanted interference is given by scalar products of independent vectors with zero mean, since the strong law of large numbers indicates that this type of interference goes to 0 as M approaches infinity. As a result, the interference can be removed. On the other hand, desired random quantities, such as received signal power, converge to non-zero scalar numbers. Beside these nice properties, simple transmission and detection schemes can be used to achieve the optimal performance [15].

We will now investigate how our random pilot detection scheme can be enhanced by massive MIMO in the presence of noise. To this end, we study z_{12} and z_{12}^E in the noisy case, when $M \to \infty$. The strong law of large numbers implies

[3]Finding the optimal expression of $r(c_B)$ is left for future work.

that the scalar product between different vectors in (6) and (9), converges to 0. Conversely, $\|\mathbf{h}_B\|^2/M$ and $\|\mathbf{h}_E\|^2/M$ converge to 1. Hence, it holds that

$$\lim_{M\to\infty} z_{12} = p_1^B (p_2^B)^* d_B^2 \tag{14}$$

and

$$\lim_{M\to\infty} z_{12}^E = p_1^B (p_2^B)^* d_B^2 + p_1^E (p_2^E)^* d_E^2. \tag{15}$$

Note that as before, in the absence of Eve, (15) is a scaled PSK symbol. Similarly to Theorem 1, we have the following theorem about the detection probability.

Theorem 2. *When $M \to \infty$, Eve can be detected with probability $1 - 2/N$.*

Proof: Write (15) as

$$\lim_{M\to\infty} z_{12}^E = p_1^B (p_2^B)^* (d_B^2 + (p_1^B)^* p_2^B p_1^E (p_2^E)^* d_E^2). \tag{16}$$

Eve will be undetected if the angle of the above product is that of a PSK symbol. Since d_B and d_E are random variables, the probability of this event is 0 if $(p_1^B)^* p_2^B p_1^E (p_2^E)^* \neq \pm 1$. Instead, if $(p_1^B)^* p_2^B p_1^E (p_2^E)^* = \pm 1$, the angle of the product is always a PSK symbol. Hence, for Eve to remain undetected, in the second time slot it must guess the pilot $p_2^E = (p_1^B)^* p_2^B p_1^E$ or $p_2^E = -(p_1^B)^* p_2^B p_1^E$, which happens with probability $2/N$. Thus, the detection probability is $1 - 2/N$. ∎

Comparing Theorems 1 and 2, we see that the probability of not detecting Eve is doubled with massive MIMO. However, as in the noiseless case, the detection probability can be made arbitrarily close to 1 by increasing the alphabet size. Moreover, this performance is guaranteed *for any noise power* and that only two training slots are needed to achieve this performance.

VI. 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. A high false-alarm probability results in pausing periods for Alice even in the absence of Eve, which decreases the average throughput to Bob. Hence, a high detection probability and a low false-alarm probability is the desired goal for detection schemes.

Figure 4 shows the detection and false-alarm probabilities versus SNR of our scheme for 8 antennas at Alice, $N = 4$ and $N = 8$ PSK, and different transmission power at Eve. Here, SNR is defined as $\mathrm{SNR} = P_B/N_0$. Moreover, we assume a large-scale fading scenario as in [13], i.e., $d_B = d_E = 1$. As expected, the detection probability increases with SNR, while the false-alarm probability decreases; in the high SNR region, we obtain the performance of Theorem 1. Increasing the alphabet size deteriorates the performance for small SNRs, but improves it for high SNRs due to Theorem 1. Interestingly, when Eve's transmit power is much larger than Bob's, our scheme provides excellent performance. The reason is that Eve "reveals" itself more, and our scheme has the ability to identify this. Hence, beside being robust to noise, it is also robust to power variations in Eve.

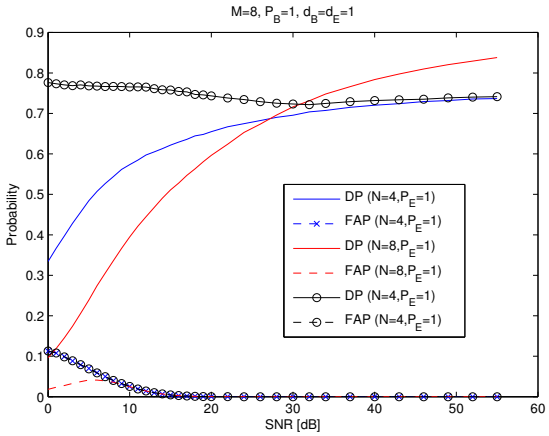

Fig. 4. Detection (DP) and false-alarm (FAP) probabilities vs. SNR. The detection probability increases with SNR, while the false-alarm probability decreases. Larger alphabets converge to a larger value, but at a slower pace than smaller alphabets. Moreover, the scheme is robust to power variations at Eve as demonstrated by the curve with circular markers.

Emulating a massive MIMO scenario, Figure 5 shows the impact of large antennas on our detection scheme. In the limit of infinite antennas, we obtain the performance given by Theorem 2. Compared to Figure 4, the performance is improved with many antennas. Note that the false-alarm probability is roughly almost 0 for all curves. Larger alphabets give better performance for high SNR, but they have a slower convergence rate. However, increasing the number of antennas to $M = 400$ gives better convergence rate, which is demonstrated by the black curve for $N = 64$ (this curve would converge slower to its limit for $M = 200$). Note that the curves are above the $1 - 2/N$ probability in Theorem 2. The reason for this is that M is still a finite number, and thus the scalar product $\mathbf{h}_B^* \mathbf{h}_E/M$ between Bob's and Eve's channels is not exactly 0. Therefore, when $\mathrm{SNR} \to \infty$, the performance is governed by Theorem 1 instead. Hence, as a result, a detection probability larger than $1 - 2/N$ can be achieved for high SNRs. As in Figure 4, increasing Eve's transmission power makes it easier to detect her presence. Note that the black curve confirms the results in Theorems 1 and 2, namely that our scheme can achieve detection probabilities arbitrarily close to 1.

VII. Conclusions and Future Work

In this paper, we presented novel detection schemes based on random pilots to combat the pilot contamination attack by a malicious user. The detection schemes require only two training slots to perform detection at the base station without any prior channel knowledge; thus only a small overhead is incurred. In the absence of noise (high SNR regime) and with large alphabet cardinality, we have revealed that our scheme achieves perfect detection. For a finite number of antennas and with the presence of additive white Gaussian noise, we also studied the detection region. Simulations results have shown that the detection scheme provides a high detection probability and low false alarm probability. The detection problem is

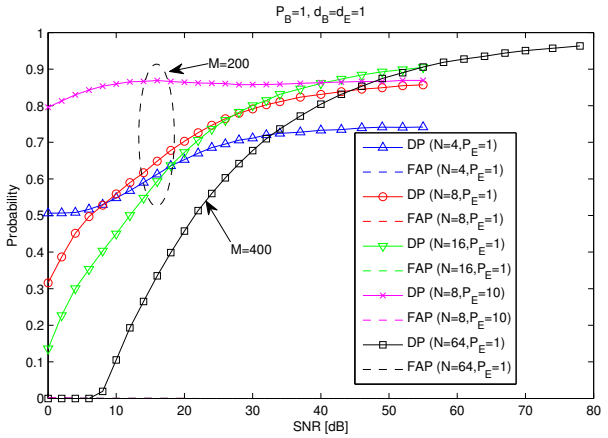

Fig. 5. Detection (DP) and false-alarm (FAP) probabilities vs. SNR for $M = 200$ and $M = 400$. The detection probability increases with SNR, while the false-alarm probability is 0. Larger alphabets have slower convergence rate, which can be improved by increasing the number of antennas at Alice. All curves converge to their limit $1 - 1/N$ predicted by Theorem 1, since the scalar product $\mathbf{h}_B^* \mathbf{h}_E / M$ is not exactly 0 for $M = 200$ and $M = 400$.

further enhanced by massive MIMO techniques and it is shown that our scheme again achieves perfect detection.

This area is not well explored. Future work includes the study of the optimal distance expressions for $r(d_B^2 \|\mathbf{h}_B\|^2)$ for the finite case, which is crucial for the performance of the detection scheme. Another important direction is to study the impact of the training phase duration on the performance. A training duration of K slots gives $\binom{K}{2}$ different scalar products, from which more information can be deduced about Eve's presence at a cost of increased overhead and complexity.

REFERENCES

[1] A. D. Wyner, "The wire-tap channel," *Bell System Tech. J.*, vol. 54, pp. 1355–1387, Oct. 1975.

[2] A. O. Hero, "Secure space-time communication" *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3235–3249, Dec. 2003.

[3] Y. Liang, H. V. Poor and S. Shamai (Shitz), "Secure communication over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2470–2492, Jun. 2008.

[4] C. Ling, L. Luzzi, and J.-C. Belfiore, "Lattice codes achieving strong secrecy over mod-Λ Gaussian channel," in *Proc. IEEE Int. Sym. Inf. Theory*, pp. 2306–2310, Cambridge, USA, Jul. 2012.

[5] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.

[6] A. L. Swindlehurst, "Fixed SINR solutions for the MIMO wiretap channel," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Process.*, pp. 2437–2440, Taiwan, Apr. 2009.

[7] L. Dong, Z. Han, A. P. Petropulu and H. V. Poor, "Improving wireless physical layer security via cooperating relays," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1875–1888, Mar. 2010.

[8] G. Zheng, Li-Chia Choo, and K. K. Wong, "Optimal cooperative jamming to enhance physical layer security using relays," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1317–1322, Mar. 2011.

[9] S. Gerbracht, C. Scheunert and E. A. Jorswieck, "Secrecy outage in MISO systems with partial channel information," *IEEE Trans. Inf. Forens. Sec.*, vol. 7, no. 2, pp. 704–716, Apr. 2012.

[10] S. Luo, J. Li and A. Petropulu, "Outage constrained secrecy rate maximization using cooperative jamming," in *Proc. 2012 IEEE Statistical Signal Process. Workshop*, Ann Arbor, MI, Aug. 2012.

[11] Y. O. Basciftci, C. E. Koksal and F. Ozguner, "To obtain or not to obtain CSI in the presence of hybrid adversary," available online http://arxiv.org/abs/1301.6449.

[12] S. M. Perlaza, A. Chorti, H. V. Poor and Z. Han, "On the tradeoffs between network state knowledge and secrecy," *Global Wireless Summit*, Atlantic City, New Jersey, USA, 24-27 Jun. 2013.

[13] X. Zhou, B. Maham, and A. Hjørungnes, "Pilot contamination for active eavesdropping," *IEEE Trans. Wireless Commun.*, pp. 903–907, Mar. 2012.

[14] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[15] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[16] A. Mukherjee and A. L. Swindlehurst, "Detecting passive eavesdroppers in the MIMO wiretap channel," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Process.*, pp. 2809–2812, Kyoto, Japan, Mar. 2012.