Pilot Spoofing Attack Detection and Countermeasure

Jitendra K. Tugnait, *Fellow, IEEE*

***Abstract*—In a time-division duplex (TDD) multiple antenna system, the channel state information (CSI) can be estimated using reverse training. A pilot spoofing (contamination) attack occurs when during the training phase, an adversary (spoofer) also sends synchronized, identical training (pilot) signal as that of the legitimate receiver. This contaminates channel estimation and alters the legitimate beamforming/precoder design, facilitating eavesdropping. A recent approach proposed superimposing a random sequence on the training sequence at the legitimate receiver and then using the minimum description length (MDL) criterion to detect pilot contamination attack. In this paper, we augment this approach with estimation of both legitimate receiver and eavesdropper channels, and secure beamforming, to mitigate the effects of pilot spoofing. We consider two cases: (i) the spoofer transmits only the pilot signal, (ii) the spoofer also adds a random sequence to its pilot, mimicking the legitimate receiver. We also employ a random matrix theory based source enumeration approach instead of MDL, for spoofing detection, leading to improved detection performance. The proposed detection and mitigation approaches are illustrated via simulations.**

***Index Terms*—Physical layer security, pilot spoofing attack, active eavesdropping, secure beamforming.**

I. Introduction

Wireless networks are vulnerable to malicious attacks aimed at disrupting their operation, due to their broadcast nature. Physical layer security methods have shown that using the properties of physical channels, it is possible to ensure confidentiality of data against eavesdropping [1], [2]. In these physical layer security methods, full or partial knowledge of the channel state information (CSI) of the legitimate system is required [2]. This knowledge is typically acquired by channel estimation during the training phase before the information signal transmission. In a time-division duplex (TDD) multiple antenna system, CSI can be acquired using reverse training.

Consider a three-node TDD multiple antenna system, consisting of a multi-antenna base station Alice, a single antenna legitimate user Bob, and a single antenna eavesdropper Eve. Alice designs its transmit beamformer based upon its channel to Bob for improved performance. In a TDD system, the downlink and uplink channels can be assumed to be reciprocal. Therefore, Alice can acquire the CSI regarding Alice-to-Bob channel via reverse training during the uplink transmission. Bob sends pilot (training) signals to Alice during the training phase of the slotted TDD system. If a publicly known protocol is used where the pilot sequences are publicly known, a malicious single-antenna terminal (eavesdropper) Eve can transmit the same pilot sequence during the training phase. If Eve attacks the channel training phase by transmitting the same pilot sequence during the training phase, synchronized with Bob's training, the CSI estimated by Alice then is a weighted sum of Bob-to-Alice and Eve-to-Alice CSIs. Consequently, the beamformer designed on this basis will lead to a significant information leakage to Eve. This is an example of a pilot spoofing/contamination attack [3]–[6].

Several types of eavesdropping have been identified and analyzed in the literature [6]. In passive eavesdropping, the eavesdropper does not transmit any signal of its own, but tries to intercept confidential communication between a legitimate transmitter-receiver pair. In active eavesdropping, the eavesdropper also transmits a signal of its own. If the intent is to disrupt the legitimate operation, active eavesdropping attack is more appropriately termed as a jamming attack [7], [8]. Such jamming attacks may occur during the training phase (pilot jamming), as in [7]–[9], and/or in the data phase, as in [7], [8]. The objective of a jamming attack is to degrade the overall legitimate system performance. Distinct from pilot jamming is the pilot spoofing or pilot contamination attack [3], [6], [10], [11], where the eavesdropper Eve sends synchronized, identical training (pilot) signal as that of the legitimate user Bob. In contrast, in a pilot jamming attack, Eve's signal is a different pilot or not noise-like signal [9]. Eve's objective in pilot spoofing is to deceive Alice into treating the Alice-to-Eve channel as Alice-to-Bob channel. This paper is concerned with pilot spoofing attack issues.

A. Related Work

There exist several approaches to the problem of pilot jamming and spoofing detection, and countermeasures to the attack. They differ in the underlying assumptions regarding the attack model, and the assumed degree of cooperation between Alice and Bob, e.g., two-way methods utilizing both uplink and downlink communication between Alice and Bob, versus methods based on uplink signals only.

1) Jamming Detection and Mitigation: Here representative examples include [7]–[9]. These approaches do not address issues involving pilot spoofing.

2) Pilot Spoofing Attack Detection: The pilot contamination attack was first noted in [3] where the focus is on enhancing eavesdropper's performance. Several approaches have been discussed in [4]–[6], [10]–[16] for detection of the attack.

a) Two-way Methods or User Cooperation: The approaches of [10], [11], [14], [15] all require two-way communication between Alice and Bob. Asymmetry of received

J.K. Tugnait is with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL 36849, USA. E-mail: tugnajk@eng.auburn.edu .

This work was supported by NSF Grant ECCS-1651133.

signals' power levels at Alice and Bob is exploited in [10] to design an attack detector. It is based on uplink reverse training from Bob to Alice as well as a subsequent downlink transmission from Alice to Bob relaying received signal power level of Bob-to-Alice transmission. A two-way training method is proposed in [14] for attack detection and secure transmission (attack mitigation), where one requires additional downlink training from Alice to Bob, not attacked by Eve. An interesting secret key arrangement protocol is proposed in [11] for attack detection and mitigation. This protocol involves several uplink and downlink transmissions between Alice and Bob in order to operate successfully. More recently, [15] proposed a two-stage uplink training method for attack detection (and secure transmission). This training method consists of splitting the training sequence into two subsequences of unequal lengths and unequal power levels from Bob, but same power level from Eve, where the knowledge of the power split is unknown to Eve, but is communicated by Alice to Bob in a downlink transmission.

b) Only Uplink Signals: The approaches proposed in [4]–[6], [12], [13], [16] for detection of the pilot spoofing attack are all based on uplink signals only. The methods discussed in [4]–[6] apply only to massive MIMO systems, relying on the fact that number of antennas at Alice is large (approaching infinity). The detection approach of [13] requires knowledge of the statistical CSI of both Bob and Eve at Alice. The training signal is self-contaminated by Bob in the method of [12], which results in signal subspace of dimension two in the presence of pilot contamination attack and of dimension one in its absence. In [12], the minimum description length (MDL) source enumeration method ([17]–[19]) based on data correlation matrix is used for estimation of the signal subspace dimension, hence, for attack detection. An additional random training phase is added in [16] after the conventional pilot training phase in order to devise a new pilot spoofing attack detector, and a secure transmission scheme. In [20], [21], a space division multiple access (SDMA) uplink was considered to allow for simultaneous transmission of training from multiple legitimate users. The model of [20], [21] also allows multiple Eves.

3) Pilot Spoofing Attack Countermeasures: Having detected a pilot spoofing attack at Alice, what approaches can be taken by Alice to secure her downlink transmission to Bob against Eve's eavesdropping? This issue has been considered in [10], [11], [14]–[16], [22], where Alice needs to obtain estimates of channels of both Bob and Eve.

a) Two-way Methods or User Cooperation: The methods of [10], [11], [14], [15] have already been discussed in Sec. I-A2a. Given the presence of attack, in the approach of [22], Alice communicates this information in downlink to Bob, who then transmits a random binary sequence in a post training phase to help devise countermeasures.

b) Only Uplink Signals: The secure transmission approach of [16] assumes that during the random training phase, Eve's power is much lower than Bob's power at Alice, otherwise Eve's channel can not be reliably estimated at Alice to allow for secure beamforming. In this paper, we allow Eve's power to be (much) higher than Bob's power at Alice, and still have secure beamforming at Alice. A preliminary conference version of this paper is [23]..

B. Contributions of This Paper

A preliminary conference version of this paper is [23], where the approach of [12] utilizing MDL criterion for attack detection, was augmented with estimation of both Bob's and Eve's channels, and secure beamforming, to mitigate the effects of pilot spoofing. In the set-up of [23], while Bob superimposes a random sequence on its training sequence, Eve does not. In this paper, we also consider the scenario where Eve also could superimpose its own random sequence on the training sequence, before spoofing Bob. The main contributions of this paper are as follows:

- Our approaches need only uplink signals at Alice in the training phase in order to detect an attack and to devise a countermeasure, unlike other existing approaches discussed in Sec. I-A, except for [16]. Unlike [16], we allow Eve's power to be (much) higher than Bob's power at Alice, and still have secure beamforming at Alice, since Bob's channel can be accurately estimated.
- We employ the random matrix theory (RMT) based source enumeration approach of [24], [25] instead of MDL, for spoofing detection, leading to more accurate detection performance, compared to our earlier results in [12], [23].
- Accurate estimation of Bob's channel in the presence of Eve's attack in uplink is crucial for secure beamforming at Alice for downlink transmission. For the case where Eve transmits only pilot in the training phase, we present a novel algorithm for estimation of Bob's and Eve's channels in Sec. III using the correlation matrices of data and projected data, where the latter is obtained by exploiting the temporal subspace properties of the pilot signal (Sec. III-B1). While various steps of the algorithm rely on existing techniques, pre-processing of data before application of these techniques, and order of the steps, are novel and non-trivial. For instance, second-order statistics-based blind estimation of Bob's channel detailed in Sec. III-B2 will not work if applied to the received signal at Alice; it has to operate on projected data which is devoid of pilot signal contributions.
- We also consider the attack model where Eve also could superimpose its own random sequence on the training sequence. We present a novel algorithm in this case for estimation of Bob's and Eve's channels in Sec. IV, using higher-order statistics of the data after removal of the contribution of pilot sequence.
- We exploit the signal subspace rank of the projected data to determine which attack model is true: Eve with pilot only, or Eve with pilot and random sequence. Thus, one does not need to know *apriori* which attack model applies. This is also a contribution of this paper.

The rest of the paper is organized as follows. In Sec. II, we present our system model, and also review some background material regarding the pilot spoofing attack detection approach of [12], and the RMT-based source enumeration approach of

[24], [25]. Two attack models are considered: (i) Eve transmits only the pilot signal, (ii) Eve also adds a random sequence to its pilot, mimicking Bob. Estimation of both Bob's and Eve's channels are discussed in Secs. III and IV for the case of pilot-only Eve transmission and pilot-and-random signal Eve transmission, respectively. In Sec. V, we discuss measures taken at Alice for her transmission to Bob, taking into account presence/absence of Eve. Simulation results are presented in Sec. VI to illustrate the proposed approaches. Since Eve's objective in pilot spoofing is to deceive Alice into treating the Alice-to-Eve channel as Alice-to-Bob channel, Bob's secrecy rate as it pertains to Alice's downlink transmission, is taken as a performance measure.

Notation: Superscripts $(.)^*$, $(.)^\top$ and $(.)^H$ represent complex conjugate, transpose and complex conjugate transpose (Hermitian) operation, respectively, on a vector/matrix. The notation $\mathbb{E}\{.\}$ denotes the expectation operation, $\mathbb{C}$ the set of complex numbers, and $\mathbf{I}_M$ an $M \times M$ identity matrix. The notation $\mathbf{x} \sim \mathcal{N}_c(\mathbf{m}, \Sigma)$ denotes a random vector $\mathbf{x}$ that is circularly symmetric complex Gaussian with mean $\mathbf{m}$ and covariance Σ. The abbreviations w.p.1 and i.p. stand for with probability one and in probability, respectively.

II. System Model and Background

We consider an MISO (multiple-input single-output) system with a multi-antenna transmitter Alice equipped with $N_r \geq 3$ antennas and a single antenna legitimate user Bob, operating in a flat Rayleigh fading environment. Alice designs its transmit beamformer for downlink transmission to Bob based on its channel to Bob. The system operates in a TDD mode where the downlink and uplink channels can be assumed to be reciprocal. If N_r is large (as in massive MIMO and similar systems [6]), it is more convenient and efficient for Bob to send a single training sequence to Alice to enable Alice to estimate the N_r sub-channels via reverse training in uplink, and then use channel reciprocity to infer Alice-to-Bob channel for downlink beamforming. If Alice were to send pilot sequences to Bob simultaneously in downlink, one would need N_r orthogonal pilot sequences, hence, at least N_r symbols long pilots. Alternatively, each Alice antenna-to-Bob sub-channel can be trained by the same pilot in downlink, sent sequentially in a time-division set-up, still needing at least N_r symbols for training. Therefore, reverse training is preferred. We assume that there exists a pilot spoofing eavesdropper Eve who transmits synchronized, identical pilot signal as that of Bob, during the training phase. Eve's objective in pilot spoofing is to deceive Alice into treating the Alice-to-Eve channel as Alice-to-Bob channel. Hence, the number of antennas at Eve must be the same as the number of antennas at Bob. Therefore, in our model, Eve also has a single antenna.

Such a system model has also been investigated in [4], [10], [12]–[16], [22]. Let $s_t(n)$, $1 \leq n \leq T$, denote the training sequence of length T time samples. Consider a flat Rayleigh fading environment with Bob-to-Alice channel denoted as $\mathbf{h}_B = \sqrt{d_B}\,\tilde{\mathbf{h}}_B \in \mathbb{C}^{N_r}$ and Eve-to-Alice channel denoted as $\mathbf{h}_E = \sqrt{d_E}\,\tilde{\mathbf{h}}_E \in \mathbb{C}^{N_r}$, where real scalars d_B and d_E include the effects of respective path losses (dependent upon path loss exponents and distance between transmitter-receiver pairs), and $\tilde{\mathbf{h}}_B \sim \mathcal{N}_c(0, \mathbf{I}_{N_r})$ and $\tilde{\mathbf{h}}_E \sim \mathcal{N}_c(0, \mathbf{I}_{N_r})$ represent small-scale fading. Let P_B and P_E denote the average training power allocated by Bob and Eve, respectively. In the absence of any transmission from Eve, the received signal at Alice during the training phase is given by

$$\mathbf{y}(n) = \sqrt{P_B}\,\mathbf{h}_B s_t(n) + \mathbf{v}(n) \tag{1}$$

where additive noise $\mathbf{v}(n) \sim \mathcal{N}_c(0, \sigma_v^2 \mathbf{I}_{N_r})$, and we normalize $T^{-1}\sum_{n=1}^{T} |s_t(n)|^2 = 1$ (e.g., take $|s_t(n)| = 1$). When Eve also transmits (Eve's pilot spoofing attack), the received signal at Alice during the training phase is

$$\mathbf{y}(n) = \left(\sqrt{P_B}\,\mathbf{h}_B + \sqrt{P_E}\,\mathbf{h}_E\right) s_t(n) + \mathbf{v}(n). \tag{2}$$

In case of Eve's attack, based on (2), Alice would estimate $\sqrt{P_B}\,\mathbf{h}_B + \sqrt{P_E}\,\mathbf{h}_E$ as Bob-to-Alice channel, instead of $\sqrt{P_B}\,\mathbf{h}_B$ based on (1).

How to detect Eve's attack based only on the knowledge of $s_t(n)$ and $\mathbf{y}(n)$, is addressed in [12], where a fraction β of the training power P_B at Bob is allocated to a scalar random sequence $s_B(n)$ (zero-mean, i.i.d., normalized to have $T^{-1}\sum_{n=1}^{T} |s_B(n)|^2 = 1$, finite alphabet: binary phase-shift keying (BPSK) or quadrature phase-shift keying (QPSK), e.g.), to be transmitted by Bob along with $s_t(n)$. That is, instead of $\sqrt{P_B}s_t(n)$, Bob transmits ($0 \leq \beta < 1, \;\; n = 1, 2, \cdots, T$)

$$\tilde{s}_B(n) = \sqrt{P_B(1-\beta)}\,s_t(n) + \sqrt{P_B\beta}\,s_B(n). \tag{3}$$

The sequence $\{s_B(n)\}$ is unknown to Alice (and to Eve) and it can not be replicated in advance as it is a random sequence generated at Bob. However, Alice knows that such $\{s_B(n)\}$ is to be expected in $\mathbf{y}(n)$. In the approach of [12], one considers the following two hypotheses, $\mathcal{H}_0$ (no attack) and $\mathcal{H}_1$ (attack present), for the received signal at Alice:

$$\begin{array}{ll} \mathcal{H}_0: & \mathbf{y}(n) = \mathbf{h}_B \tilde{s}_B(n) + \mathbf{v}(n) \\ \mathcal{H}_1: & \mathbf{y}(n) = \mathbf{h}_B \tilde{s}_B(n) + \sqrt{P_E}\,\mathbf{h}_E s_t(n) + \mathbf{v}(n). \end{array} \tag{4}$$

We also consider another attack model where we allow Eve too to add a scalar random sequence $s_E(n)$ (clearly independent of $s_B(n)$) to $s_t(n)$ at Eve's transmitter. This results in the following set-up:

$$\begin{array}{ll} \mathcal{H}_0: & \mathbf{y}(n) = \mathbf{h}_B \tilde{s}_B(n) + \mathbf{v}(n) \\ \mathcal{H}_1: & \mathbf{y}(n) = \mathbf{h}_B \tilde{s}_B(n) + \mathbf{h}_E \tilde{s}_E(n) + \mathbf{v}(n) \end{array} \tag{5}$$

where

$$\tilde{s}_E(n) = \sqrt{P_E(1-\beta_2)}\,s_t(n) + \sqrt{P_E\beta_2}\,s_E(n), \tag{6}$$

and Eve allocates a fraction β_2 of its transmit power to $\{s_E(n)\}$. We assume that $\{s_E(n)\}$ is similar in nature to $\{s_B(n)\}$, i.e., it is zero-mean, i.i.d., finite alphabet (but unknown to Alice), and normalized to have $T^{-1}\sum_{n=1}^{T} |s_E(n)|^2 = 1$. That is, Eve spoofs Bob more closely than in [12].

A. Attack Detection Approach of [12]

It is based on the model (4). Define the correlation matrix of measurements as ($i = 0, 1$)

$$\mathbf{R}_{y,i} = T^{-1} \sum_{n=1}^{T} \mathbb{E}\left\{\mathbf{y}(n)\mathbf{y}^H(n) \,\middle|\, \mathcal{H}_i\right\} \tag{7}$$

and the correlation matrix of source signals as ($i = 0, 1$)

$$\mathbf{R}_{s,i} = T^{-1} \sum_{n=1}^{T} \mathbb{E}\left\{[\mathbf{y}(n) - \mathbf{v}(n)][\mathbf{y}(n) - \mathbf{v}(n)]^H \,\middle|\, \mathcal{H}_i\right\}. \tag{8}$$

Then we have

$$\mathbf{R}_{y,i} = \mathbf{R}_{s,i} + \sigma_v^2 \mathbf{I}_{N_r}, \quad i = 0, 1. \tag{9}$$

It is shown in [12] that $\text{rank}(\mathbf{R}_{s,0}) = 1$ and $\text{rank}(\mathbf{R}_{s,1}) = 2$. Thus, introduction of $\{s_B(n)\}$ by Bob leads to signal subspace of rank 2 in the presence of Eve's attack. If $\beta = 0$, then $\text{rank}(\mathbf{R}_{s,1}) = 1$. Define the sample correlation matrix as

$$\widehat{\mathbf{R}}_y = T^{-1} \sum_{n=1}^{T} \mathbf{y}(n)\mathbf{y}^H(n). \tag{10}$$

Let the ordered eigenvalues of $\widehat{\mathbf{R}}_y$ be denoted by $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{N_r}$. The MDL estimator [17] $\widehat{d}_{MDL}$ of the signal subspace dimension based on the eigenvalues λ_is, is exploited in [12] to detect spoofing attack; for details, see [12]. If $\widehat{d}_{MDL} = 1$, there is no spoofing pilot attack, and if $\widehat{d}_{MDL} > 1$, we have a spoofing pilot attack. Attack mitigation (countermeasure) is not addressed in [12]. Note that if Eve also adds a random sequence to its pilot, $\text{rank}(\mathbf{R}_{s,1}) = 2$. The attack will still be detected, i.e., the approach of [12] also applies to attack model (5).

B. RMT Based Rank Determination [24], [25]

The RMT-based approach of [24], [25] is based on the distribution of the eigenvalues of $\widehat{\mathbf{R}}_y$ when $\mathbf{y}(n) = \mathbf{v}(n)$ (noise eigenvalues). Let $\widehat{\mathbf{R}}_v = T^{-1} \sum_{n=1}^{T} \mathbf{v}(n)\mathbf{v}^H(n)$, with ordered eigenvalues $\ell_1 \geq \ell_2 \geq \cdots \geq \ell_{N_r}$. As $N_r, T \to \infty$ with $N_r/T \to c \geq 0$, the largest eigenvalue ℓ_1 of $\widehat{\mathbf{R}}_v$ is distributed with a Tracy-Widom distributed of order 2 [25], i.e.,

$$\lim_{\substack{N_r,T\to\infty \\ N_r/T\to c\geq 0}} P\left\{\ell_1 < \sigma_v^2\left(\mu_{T,N_r} + z\sigma_{T,N_r}\right)\right\} = F_{TW,2}(z) \tag{11}$$

where $F_{TW,2}(z)$ denotes the Tracy-Widom probability distribution function of order 2, and by [26],

$$\begin{aligned} \mu_{T,N_r} =& \frac{1}{T}\left(\sqrt{T+0.5} + \sqrt{N_r+0.5}\right)^2, \\ \sigma_{T,N_r} =& \sqrt{\frac{\mu_{T,N_r}}{T}}\left(\frac{1}{\sqrt{T+0.5}} + \frac{1}{\sqrt{N_r+0.5}}\right)^{1/3}. \end{aligned} \tag{12}$$

When signals with correlation matrix of rank d are added to $\mathbf{v}(n)$, the $(d+1)$st eigenvalue of $\widehat{\mathbf{R}}_y$ follows a Tracy-Widom distribution, as in (11), except that we replace N_r with $N_r - d$; for details, see [25].

The RMT-based signal subspace rank estimator $\widehat{d}_{RMT}$ of [24], [25] is given by

$$\widehat{d}_{RMT} = \arg\min_k \left\{\lambda_k < \hat{\sigma_v}^2(k)\left(\mu_{T,N_r-k} + z(P_{fa})\sigma_{T,N_r}\right)\right\} - 1 \tag{13}$$

where $k \in \{1, 2, \cdots, \min(N_r, T) - 1\}$, $z(P_{fa})$ is such that $F_{TW,2}(z(P_{fa})) = P_{fa}$, and P_{fa} is the probability of false alarm for the hypothesis testing problem where signal subspace rank $\leq k-1$ is the null hypothesis $\mathcal{H}_0$ and signal subspace rank $\geq k$ is the alternative $\mathcal{H}_1$. Instead of using the maximum likelihood estimator $\hat{\sigma_v}^2(k) = \frac{1}{N_r-k}\sum_{i=k+1}^{N_r}\lambda_i$ for noise variance (which has a negative bias [25]), [25] develops an alternative iterative estimator $\hat{\sigma_v}^2(k)$ of σ_v^2 which is more accurate. A MATLAB code of the algorithm of [24], [25] is available at http://www.wisdom.weizmann.ac.il/~nadler/Rank_Estimation. Since in this paper, we are interested only in the fact whether $\widehat{d}_{RMT} = 1$ or $\widehat{d}_{RMT} > 1$, in (13), we search for the minimum over a smaller set $k \in \{1, 2\}$. As discussed in [24], [25], the RMT-based approach can detect weaker signals compared to the MDL detector.

III. Channel Estimation Under Attack Model (4)

In this section, we discuss several variations of channel estimation assuming that the attack model (4) is true. If no attack is indicated, Alice uses an iterative channel estimation approach to estimate Bob's channel, as discussed in Sec. III-A. It turns out that independent of attack detection, Alice can estimate Bob's channel up to a complex constant. This is discussed in Sec. III-B. If the MDL or the RMT method indicates presence of attack, Alice can jointly estimate the channels to Bob and Eve. These estimated channels underlie the countermeasures taken at Alice for her transmission to Bob, discussed in Sec. V.

A. No Attack

If the attack detector indicates absence of any attack, Alice proceeds to initially estimate the channel using (4) under $\mathcal{H}_0$, knowledge of $\{s_t(n)\}$ and the least-squares method. This approach treats $\{s_B(n)\}$ as interference. An obvious solution, as discussed in [12], is to perform iterative channel estimation via a linear minimum mean-square error (MMSE) equalizer to estimate and decode (quantize) self-contamination $s_B(n)$, and then use the decoded $s_B(n)$ in conjunction with $s_t(n)$ as pseudo-training. For details, see [12, Sec. IV].

B. Blind Estimation of Bob's Channel

Here we exploit temporal subspace properties of the pilot signal to null out the contribution of the pilot, from both Bob and Eve (if Eve's signal is present), to the signal received at Alice. However, the contribution of random sequence $s_B(n)$ remains, and is exploited to estimate Alice-to-Bob channel, up to a complex constant. This method applies whether or not a pilot spoofing attack is present, and whether or not the attack detector (RMT or MDL) detects a spoofing attack.

1) Projection Orthogonal to Training: The projection approach discussed here was first proposed in the preliminary conference version [23]. It was later used in [21] for detection and identification of spoofed pilots in TDD/SDMA systems where multiple Bobs and Eves are allowed. However, [21] does not address the issue of estimation of Bob's channel when Bob is under a pilot spoofing attack.

Stack P consecutive samples of ℓth component $y_\ell(n)$ of $\mathbf{y}(n)$ into a column:

$$\underbrace{y_\ell(1)\ \cdots\ y_\ell(P)}_{\mathbf{y}^\ell(1)}\ \underbrace{y_\ell(P+1)\ \cdots\ y_\ell(2P)}_{\mathbf{y}^\ell(2)}\cdots \tag{14}$$

Define $\mathbf{v}^\ell(m)$ from $v_\ell(n)$, the ℓth component $\mathbf{v}(n)$, in a similar fashion. Let

$$\check{\mathbf{s}}_t = [s_t(1)\ s_t(2)\ \cdots\ s_t(P)]^\top, \tag{15}$$

$$\check{\mathbf{s}}_B(m) = [s_B(1+(m-1)P)\ \cdots\ s_B(P+(m-1)P)]^\top. \tag{16}$$

Then, in the presence of self-contamination and eavesdropper, we have

$$\begin{aligned}\mathbf{y}^\ell(m) = &\left(\sqrt{P_B(1-\beta)}\,h_{B,\ell} + \sqrt{P_E}\,h_{E,\ell}\right)\check{\mathbf{s}}_t \\ &+ \sqrt{P_B\beta}\,h_{B,\ell}\check{\mathbf{s}}_B(m) + \mathbf{v}^\ell(m)\end{aligned} \tag{17}$$

where $h_{B,\ell}$ is the ℓth component of $\mathbf{h}_B$, and similarly for $h_{E,\ell}$. Let $\mathcal{P}^\perp_{\check{\mathbf{s}}_t}$ denote the projection orthogonal to the subspace spanned by $\check{\mathbf{s}}_t$, given by $\mathcal{P}^\perp_{\check{\mathbf{s}}_t} = \mathbf{I}_P - P^{-1}\check{\mathbf{s}}_t\check{\mathbf{s}}_t^H \in \mathbb{C}^{P\times P}$, where we have used $\check{\mathbf{s}}_t^H\check{\mathbf{s}}_t = P$. Then $\mathcal{P}^\perp_{\check{\mathbf{s}}_t}\mathbf{y}^\ell(m)$ has no contribution from training $s_t(n)$. "Reshape" $\mathcal{P}^\perp_{\check{\mathbf{s}}_t}\mathbf{y}^\ell(m)$ into a row vector along time and put all components ℓs together. Then the so "projected" $\mathbf{y}(n)$ lacks $s_t(n)$, but has the effect of $\mathbf{h}_B$ and $s_B(n)$, which can be used to estimate $\mathbf{h}_B$ up to a scale factor via eigen-decomposition. We elaborate on this approach in what follows.

The eigenvalue decomposition (EVD) of Hermitian, nonnegative definite matrix $\mathcal{P}^\perp_{\check{\mathbf{s}}_t}$ of rank $P-1$, can be expressed as

$$\mathcal{P}^\perp_{\check{\mathbf{s}}_t} = \mathbf{U}_1\Sigma_1\mathbf{U}_1^H, \quad \mathbf{U}_1 \in \mathbb{C}^{P\times(P-1)}, \tag{18}$$

where Σ_1 is diagonal with positive eigenvalues along its diagonal. Consider the reduced dimension vectors (of dimension $P-1$), given by

$$\begin{aligned}&\mathbf{v}^{\ell r}(m) := \mathbf{U}_1^H\mathbf{v}^\ell(m),\ \ \mathbf{y}^{\ell r}(m) := \mathbf{U}_1^H\mathbf{y}^\ell(m),\\ &\check{\mathbf{s}}^r_B(m) := \mathbf{U}_1^H\check{\mathbf{s}}_B(m).\end{aligned} \tag{19}$$

Then $\mathbb{E}\{\mathbf{v}^{\ell r}(m)(\mathbf{v}^{\ell r}(m))^H\} = \sigma_v^2\mathbf{I}_{P-1}$, and $\mathbf{v}^{\ell r}(m_1)$ and $\mathbf{v}^{\ell r}(m_2)$ are independent for $m_1 \neq m_2$; see also [21, Sec. III]. Since $\mathcal{P}^\perp_{\check{\mathbf{s}}_t}\check{\mathbf{s}}_t = 0$ implies $\mathbf{U}_1^H\check{\mathbf{s}}_t = 0$, we have, for $m = 1, 2, \cdots, T/P$,

$$\mathbf{y}^{\ell r}(m) = \sqrt{P_B\beta}\,h_{B,\ell}\check{\mathbf{s}}^r_B(m) + \mathbf{v}^{\ell r}(m). \tag{20}$$

Now reshape $\mathbf{y}^{\ell r}(m)$, $m = 1, \cdots, T/P$, with T/P an integer, into a row of scalars $\tilde{y}_\ell(n)$, $n = 1, 2, \cdots, (T/P)(P-1)$, using the correspondence

$$\underbrace{\tilde{y}_\ell(1)\ \cdots\ \tilde{y}_\ell(P-1)}_{\mathbf{y}^{\ell r}(1)}\ \underbrace{\tilde{y}_\ell(P)\ \cdots\ \tilde{y}_\ell(2(P-1))}_{\mathbf{y}^{\ell r}(2)}\cdots \tag{21}$$

Similarly define $\tilde{v}_\ell(n)$ from $\mathbf{v}^{\ell r}(m)$, $m = 1, \cdots, T/P$, and similarly construct $\tilde{s}^r_B(n)$ from $\check{\mathbf{s}}^r_B(m)$. Then $\tilde{\mathbf{y}}(n) \in \mathbb{C}^{N_r}$, with ℓth component $\tilde{y}_\ell(n)$, satisfies

$$\tilde{\mathbf{y}}(n) = \sqrt{P_B\beta}\,\mathbf{h}_B\tilde{s}^r_B(n) + \tilde{\mathbf{v}}(n). \tag{22}$$

In the above model, $\{\tilde{\mathbf{v}}(n)\}$ is i.i.d. zero-mean complex Gaussian with covariance $\sigma_v^2\mathbf{I}_{N_r}$, and similarly $\tilde{s}_B(n)$ is uncorrelated zero-mean sequence with $\mathbb{E}\{|\tilde{s}_B(n)|^2\}$ not a function of n (follows just as the properties of $\tilde{\mathbf{v}}(n)$).

2) Estimation of Bob's Channel: Consider (22) with $n = 1, 2, \cdots, n_b(P-1)$, where $n_b = T/P$= an integer. Then, as in (7), with $T' = n_b(P-1)$,

$$\mathbf{R}_{\tilde{y}} = \frac{1}{T'}\sum_{n=1}^{T'}\mathbb{E}\{\tilde{\mathbf{y}}(n)\tilde{\mathbf{y}}^H(n)\} = \beta P_B\mathbf{h}_B\mathbf{h}_B^H + \sigma_v^2\mathbf{I}_{N_r} \tag{23}$$

where $\mathbb{E}\{|s_B(n)|^2\} = 1 = \mathbb{E}\{|\tilde{s}_B(n)|^2\}$. It follows that $\mathbf{h}_B/\|\mathbf{h}_B\|$ is a unit-norm eigenvector of $\mathbf{R}_{\tilde{y}}$ with eigenvalue $(\beta P_B\|\mathbf{h}_B\|^2 + \sigma_v^2)$, all other eigenvectors have eigenvalues σ_v^2. Hence, we estimate $\mathbf{h}_B$ up to a complex constant as the unit norm eigenvector $\hat{\mathbf{v}}_1$ corresponding to the largest eigenvalue of $\widehat{\mathbf{R}}_{\tilde{y}}$,

$$\widehat{\mathbf{R}}_{\tilde{y}} = \frac{1}{T'}\sum_{n=1}^{T'}\tilde{\mathbf{y}}(n)\tilde{\mathbf{y}}^H(n). \tag{24}$$

The above approach can be thought of as applying the EVD-based channel estimation method of [27], proposed for multicell, multi-user massive MIMO systems, but applicable here for our problem even though N_r can be small, since we have just one user Bob in (22).

Since $\mathbf{h}_B \approx c\hat{\mathbf{v}}_1$ for some complex c, we pick c to minimize

$$\frac{1}{T}\sum_{n=1}^{T}\|\mathbf{y}(n) - c\hat{\mathbf{v}}_1\sqrt{(1-\beta)P_B}s_t(n)\|^2, \tag{25}$$

leading to the solution

$$\hat{c} = \frac{1}{\sqrt{(1-\beta)P_B}T}\sum_{n=1}^{T}(\hat{\mathbf{v}}_1^H\mathbf{y}(n))s_t^*(n). \tag{26}$$

Then we have the estimate of $\mathbf{h}_B$ as

$$\hat{\mathbf{h}}_B = \hat{c}\hat{\mathbf{v}}_1. \tag{27}$$

Since $\lim_{T\to\infty}\widehat{\mathbf{R}}_{\tilde{y}} = \mathbf{R}_{\tilde{y}}$ i.p. [18], we have $\lim_{T\to\infty}\hat{\mathbf{v}}_1 = \mathbf{v}_1$ i.p. [18], and for some θ,

$$\mathbf{v}_1 = e^{j\theta}\mathbf{h}_B/\|\mathbf{h}_B\|. \tag{28}$$

Note that the eigenvector is unit norm, and $e^{j\theta}$ reflects the complex scalar, unit modulus ambiguity that remains in determining the eigenvector. Using (4) under $\mathcal{H}_1$, (26) and (27), we have

$$\lim_{T\to\infty}\hat{c} = \mathbf{v}_1^H\left(\mathbf{h}_B + \sqrt{P_E/(P_B(1-\beta))}\,\mathbf{h}_E\right)\ \text{ i.p.} \tag{29}$$

It follows from (26)-(29) that, i.p.,

$$\lim_{T\to\infty}\hat{\mathbf{h}}_B = \bar{\mathbf{h}}_B = \left(1 + \sqrt{P_E/(P_B(1-\beta))}\,\mathbf{h}_B^H\mathbf{h}_E/\|\mathbf{h}_B\|^2\right)\mathbf{h}_B\,. \tag{30}$$

Thus, while the direction of $\bar{\mathbf{h}}_B$ is correct, its complex scaling is biased. As $N_r \to \infty$, $\mathbf{h}_B^H\mathbf{h}_E/\|\mathbf{h}_B\|^2 =$

$(1/N_r)\mathbf{h}_B^H\mathbf{h}_E/((1/N_r)\mathbf{h}_B^H\mathbf{h}_B) \to 0$ w.p.1 since $\tilde{\mathbf{h}}_B \sim \mathcal{N}_c(0,\mathbf{I}_{N_r})$, $\mathbf{h}_E \sim \mathcal{N}_c(0,\mathbf{I}_{N_r})$, and $\tilde{\mathbf{h}}_B$ and $\tilde{\mathbf{h}}_E$ are independent.

C. Under Attack

If the MDL or the RMT method indicates presence of attack, Alice can jointly estimate the channels to Bob and Eve, as discussed in this section. To this end, we strive for a more accurate estimate of $\mathbf{h}_B$ compared to that in Sec. III-B2, utilizing the finite alphabet property of $s_B(n)$.

Under $\mathcal{H}_1$, we have

$$\mathbf{y}(n) = \left(\sqrt{P_B(1-\beta)}\,\mathbf{h}_B + \sqrt{P_E}\,\mathbf{h}_E\right) s_t(n) + \sqrt{P_B\beta}\,\mathbf{h}_B s_B(n) + \mathbf{v}(n). \tag{31}$$

We estimate the composite channel

$$\mathbf{h}_c := \sqrt{P_B(1-\beta)}\,\mathbf{h}_B + \sqrt{P_E}\,\mathbf{h}_E \tag{32}$$

using the training sequence $s_t(n)$ and least-squares, as

$$\hat{\mathbf{h}}_c = \frac{1}{T}\sum_{n=1}^{T}\mathbf{y}(n)s_t^*(n). \tag{33}$$

This is an unbiased estimator of $\mathbf{h}_c$ since $\mathbb{E}\{\hat{\mathbf{h}}_c\} = \mathbf{h}_c$. Furthermore, straightforward calculations show that

$$\mathbb{E}\{\|\hat{\mathbf{h}}_c - \mathbb{E}\{\hat{\mathbf{h}}_c\}\|^2\} = \frac{1}{T}\left[\beta P_B\|\mathbf{h}_B\|^2 + \sigma_v^2 N_r\right]. \tag{34}$$

Therefore, we have $\lim_{T\to\infty}\hat{\mathbf{h}}_c = \mathbf{h}_c$ i.p.

We first estimate $s_B(n)$, up to a rotation ambiguity, using its finite alphabet property. This property does not hold for $\tilde{s}_B(n)$ in (22), which precludes use of (22) and the approach of Sec. III-B. Let χ denote the signal symbol constellation, and let Θ be the set of rotation angles which leave χ invariant, i.e., if $\chi = \{s_m,\ m = 1,2,\cdots,M\}$ and $\Theta = \{\theta_\ell,\ \ell = 1,2,\cdots,L\}$, then $\chi = \{e^{j\theta_\ell}s_m,\ m = 1,2,\cdots,M\}$ for any $\theta_\ell \in \Theta$. For example, let χ correspond to a QPSK constellation with $M = 4$ and $\chi = \{1,\ j,\ -1,\ -j\}$. Then the corresponding $\Theta = \{(\ell-1)\pi/2,\ \ell = 1,2,3,4\}$ with $L = 4$. Alice know the sets χ and Θ.

Define

$$\begin{aligned}\check{\mathbf{y}}(n) =& \mathbf{y}(n) - \hat{\mathbf{h}}_c s_t(n)\\ =& \sqrt{P_B\beta}\,\mathbf{h}_B s_B(n) + (\mathbf{h}_c - \hat{\mathbf{h}}_c)s_t(n) + \mathbf{v}(n) &(35)\\ \approx& \sqrt{P_B\beta}\,\mathbf{h}_B s_B(n) + \mathbf{v}(n). &(36)\end{aligned}$$

With $\hat{\mathbf{v}}_1$ as in Sec. III-B1, define

$$\begin{aligned}x(n) =& \hat{\mathbf{v}}_1^H\check{\mathbf{y}}(n) \approx \sqrt{P_B\beta}\,\hat{\mathbf{v}}_1^H\mathbf{h}_B s_B(n) + \hat{\mathbf{v}}_1^H\mathbf{v}(n) &(37)\\ =& c_x s_B(n) + v_x(n) &(38)\end{aligned}$$

where $v_x(n) = \hat{\mathbf{v}}_1^H\mathbf{v}(n) \sim \mathcal{N}_c(0,\sigma_v^2)$ and

$$c_x = \sqrt{P_B\beta}\,\hat{\mathbf{v}}_1^H\mathbf{h}_B \approx \sqrt{P_B\beta}\mathbf{v}_1^H\mathbf{h}_B = e^{j\theta}\underbrace{\sqrt{P_B\beta}\|\mathbf{h}_B\|}_{=\alpha}, \tag{39}$$

provided $\hat{\mathbf{v}}_1$ is close to $\mathbf{v}_1$, as defined in (28).

Using (38) and (39), one can estimate $s_B(n)$, up to a rotation ambiguity. In the following we will assume a QPSK constellation, so that in χ, $M = 4$ and $|s_m|$=1 for every $m = 1,2,3,4$. We estimate θ in (39), using (38) and assuming QPSK signals, as

$$\hat{\theta} = \arg\min_{0\le\phi\le(\pi/2)}\ r_1(\phi) \tag{40}$$

where

$$r_1(\phi) = \frac{1}{T}\sum_{n=1}^{T}\left\{\min_{\psi\in\{0,\frac{\pi}{2},\pi,\frac{3\pi}{2}\}}\left|\frac{e^{-j\phi}\,x(n)}{|x(n)|} - e^{j\psi}\right|\right\}.$$

In $r_1(\phi)$ above, we first compensate for unknown θ with $-\phi$, then normalize $e^{-j\phi}\,x(n)$ to unit modulus (the modulus of QPSK symbol constellation) which also eliminates the need to know α, and for each time n, pick the symbol in the constellation closest to $e^{-j\phi}\,x(n)/|x(n)|$ as the estimate of $s_B(n)$. Optimization in (40) can be carried out exhaustively over a grid of ϕ values in the interval $[0,\pi/2]$. Because of the rotation ambiguity in QPSK constellation, in (40), search is confined to the interval $[0,\pi/2]$, instead of $[0,2\pi)$.

With optimized $\hat{\theta}$ from (40), we estimate $s_B(n)$ as

$$\hat{s}_{Br}(n) = e^{-j\hat{\psi}(n)} \tag{41}$$

where

$$\hat{\psi}(n) = \min_{\psi\in\{0,\frac{\pi}{2},\pi,\frac{3\pi}{2}\}}\left|\frac{e^{-j\hat{\theta}}\,x(n)}{|x(n)|} - e^{j\psi}\right|.$$

In the absence of any noise, the estimated symbols $\hat{s}_{Br}(n)$ may differ from true $s_B(n)$ due to rotation ambiguity (symmetry) of χ, i.e., for QPSK constellation, $\hat{s}_{Br}(n) = e^{j\theta}s_B(n)$ for some $\theta \in \Theta$ and $n = 1,2,\cdots,T$. In the presence of noise, we additionally have symbol decoding errors.

Our next step is to resolve this rotation ambiguity, and then use the decoded symbols as pseudo-pilots, in conjunction with (36), to re-estimate $\mathbf{h}_B$ without the rotation ambiguity. Consider (36) and with $\sqrt{\beta}\hat{s}_{Br}(n)$ as the training sequence, estimate $\sqrt{P_B}\mathbf{h}_B$ as $\hat{\mathbf{h}}_{Br}$ up to a rotation ambiguity, as

$$\hat{\mathbf{h}}_{Br} = \frac{1}{\sqrt{\beta}\,T}\sum_{n=1}^{T}\check{\mathbf{y}}(n)\hat{s}_{Br}^*(n) \tag{42}$$

where we have used the fact that $|s_{Br}(n)| = 1$. We now look for an estimate $\hat{\mathbf{h}}_B = e^{j\bar{\theta}}\hat{\mathbf{h}}_{Br}$ for $\bar{\theta} \in \Theta$, which resolves the rotation ambiguity $\hat{\mathbf{h}}_{Br}$. To this end, we reconstruct Bob's contribution to $\mathbf{y}(n)$ at Alice, and cancel it from $\mathbf{y}(n)$ to leave only Eve's contribution and noise. Consider

$$\hat{\mathbf{y}}_{B,\bar{\theta}}(n) = \sqrt{1-\beta}e^{j\bar{\theta}}\hat{\mathbf{h}}_{Br}s_t(n) + \sqrt{\beta}\hat{\mathbf{h}}_{Br}\hat{s}_{Br}(n) \tag{43}$$

which is an attempt to reconstruct the contribution of Bob's signal $\sqrt{P_B(1-\beta)}\,s_t(n) + \sqrt{P_B\beta}\,s_B(n)$ to Alice's received signal $\mathbf{y}(n)$. Note that we include $\sqrt{P_B}$ in $\hat{\mathbf{h}}_{Br}$. We can not resolve the rotation ambiguity by using $\hat{s}_{Br}(n)$ alone, since the latter is the cause of this ambiguity, and furthermore, the ambiguity $\bar{\theta}$ is inherent in $\hat{s}_{Br}(n)$. However, since true training $s_t(n)$ is known, it suffers from no such uncertainty. Let

$$\mathbf{y}_{d,\bar{\theta}}(n) = \mathbf{y}(n) - \hat{\mathbf{y}}_{B,\bar{\theta}}(n) \tag{44}$$

where $\mathbf{y}(n)$ follows (4) under $\mathcal{H}_1$. When $\bar{\theta} \in \Theta$ takes its true value, we have $\mathbf{y}_{d,\bar{\theta}}(n) \approx \sqrt{P_E}\,\mathbf{h}_E s_t(n) + \mathbf{v}(n)$, otherwise it

has additional contributions due to $\mathbf{h}_B$, $\hat{\mathbf{h}}_{Br}$, $s_B(n)$, $s_t(n)$, and $\hat{s}_{Br}(n)$. Consider

$$\hat{\bar{\theta}} = \arg\min_{\bar{\theta}\in\Theta} r_2(\bar{\theta}) \text{ where } r_2(\bar{\theta}) = \frac{1}{T}\sum_{n=1}^{T} |\mathbf{y}_{d,\bar{\theta}}(n)|^2 . \quad (45)$$

When $\bar{\theta}$ is not equal to its true value, we have

$$\begin{aligned}\mathbf{y}_{d,\bar{\theta}}(n) =& \sqrt{1-\beta}\left(\sqrt{P_B}\mathbf{h}_B - e^{j\bar{\theta}}\hat{\mathbf{h}}_{Br}\right) s_t(n) \\ &+ \sqrt{\beta}\left(\sqrt{P_B}\mathbf{h}_B s_B(n) - \hat{\mathbf{h}}_{Br}\hat{s}_{Br}(n)\right) \\ &+ \sqrt{P_E}\,\mathbf{h}_E s_t(n) + \mathbf{v}(n) \quad (46)\\ \approx& \sqrt{1-\beta}\left(\sqrt{P_B}\mathbf{h}_B - e^{j\bar{\theta}}\hat{\mathbf{h}}_{Br}\right) s_t(n) \\ &+ \sqrt{P_E}\,\mathbf{h}_E s_t(n) + \mathbf{v}(n). \quad (47)\end{aligned}$$

When $\bar{\theta}$ is equal to its true value, using $\hat{\mathbf{h}}_B = e^{j\bar{\theta}}\hat{\mathbf{h}}_{Br} \approx \mathbf{h}_B$, we have

$$\mathbf{y}_{d,\bar{\theta}}(n) \approx \sqrt{P_E}\,\mathbf{h}_E s_t(n) + \mathbf{v}(n). \quad (48)$$

Thus, with $\bar{\theta}_0$ denoting the true value of $\bar{\theta}$, one expects $r_2(\bar{\theta}_0) < r_2(\bar{\theta})$ for $\bar{\theta} \neq \bar{\theta}_0$, $\bar{\theta}$, $\bar{\theta}_0 \in \Theta$, resulting in $\hat{\bar{\theta}} = \bar{\theta}_0$. Then we estimate $\sqrt{P_B}\mathbf{h}_B$ without any rotation ambiguity as

$$\hat{\mathbf{h}}_B = e^{j\hat{\bar{\theta}}}\,\hat{\mathbf{h}}_{Br}. \quad (49)$$

Since $\mathbf{y}_{d,\hat{\bar{\theta}}}(n) \approx \sqrt{P_E}\,\mathbf{h}_E s_t(n) + \mathbf{v}(n)$, a least-squares estimate $\hat{\mathbf{h}}_E$ of $\sqrt{P_E}\mathbf{h}_E$ is given by

$$\hat{\mathbf{h}}_E = \frac{1}{T}\sum_{n=1}^{T} \mathbf{y}_{d,\hat{\bar{\theta}}}(n)\hat{s}_t^*(n). \quad (50)$$

D. Summary of Solutions Under Attack Model (4)

We have discussed three different approaches to channel estimation under the attack model (4), which we now summarize. In the following, we assume that, if needed, the RMT-based spoofing attack detector is used.

1) Attack Not Detected: If the attack detector declares no attack, we can only estimate Bob's channel. If $\hat{d}_{RMT} = 1$, estimate Bob's channel via the method of [12, Sec. IV], and design Alice's beamformer as in (70) of Sec. V, discussed later. Estimation of Bob's channel utilizes both training $s_t(n)$ and random sequence $s_B(n)$ in an iterative procedure.

2) Attack Ignored – Blind Channel Estimation: Here regardless of the outcome of the spoofing attack detector, we only estimate Bob's channel, using the method of Sec. III-B. Generate the projected sequence $\tilde{\mathbf{y}}(n)$ as specified in Sec. III-B1. Then Bob's channel is estimated as in (27), and it equals a scaled version of the eigenvector of the sample correlation matrix of $\tilde{\mathbf{y}}(n)$. For secure transmission from Alice to Bob, design Alice's beamformer as in (70) of Sec. V, discussed later.

3) Attack Detected: If $\hat{d}_{RMT} > 1$, there is a spoofing attack. Now Alice estimates both Bob's and Eve's channels, following the procedure outlined in Sec. III-C.

IV. Channel Estimation Under Attack Model (5)

Now we consider the case where attack model (5) could be true. If (5) and hypothesis $\mathcal{H}_1$ are true, then (22) becomes

$$\tilde{\mathbf{y}}(n) = \sqrt{P_B\beta}\,\mathbf{h}_B\tilde{s}_B(n) + \sqrt{P_E\beta_2}\,\mathbf{h}_E\tilde{s}_E(n) + \tilde{\mathbf{v}}(n) \quad (51)$$

where $\tilde{s}_E(n)$ is defined in (6). While the signal subspace of (22) is of rank 1, the signal subspace of (51) is of rank 2. The RMT or the MDL criterion applied to $\{\tilde{\mathbf{y}}(n)\}$ will reveal its signal subspace rank. If this rank is 1, the approach of Sec. III-C for channel estimation suffices. If this rank is 2, then (5) is true, and the approach of Sec. III-C will fail. We now consider how to estimate channels of Bob and Eve in the latter case. The signal in (51) represents a mixture of two non-Gaussian signals in white Gaussian noise, but the non-Gaussian signals are not necessarily stationary (although they are wide-sense stationary), hence, the standard approaches for unmixing using higher-order statistics (e.g., kurtosis) [28]–[31] do not apply to (51). Therefore, we follow a different approach.

Under model (5) and hypothesis $\mathcal{H}_1$, let

$$\mathbf{h}_{c2} := \sqrt{P_B(1-\beta)}\,\mathbf{h}_B + \sqrt{P_E(1-\beta_2)}\,\mathbf{h}_E. \quad (52)$$

Let us estimate $\mathbf{h}_{c2}$ using the training sequence $s_t(n)$ and least-squares, as

$$\hat{\mathbf{h}}_{c2} = \frac{1}{T}\sum_{n=1}^{T}\mathbf{y}(n)s_t^*(n). \quad (53)$$

This is an unbiased estimator of $\mathbf{h}_{c2}$ since $\mathbb{E}\{\hat{\mathbf{h}}_{c2}\} = \mathbf{h}_{c2}$. Also, straightforward calculations show that

$$\mathbb{E}\{\|\hat{\mathbf{h}}_{c2} - \mathbb{E}\{\hat{\mathbf{h}}_{c2}\}\|^2\} = \frac{\left[\beta P_B\|\mathbf{h}_B\|^2 + \beta_2 P_E\|\mathbf{h}_E\|^2 + \sigma_v^2 N_r\right]}{T}. \quad (54)$$

Therefore, we have $\lim_{T\to\infty}\hat{\mathbf{h}}_{c2} = \mathbf{h}_{c2}$ i.p. Define

$$\begin{aligned}\check{\mathbf{y}}_2(n) =& \mathbf{y}(n) - \hat{\mathbf{h}}_{c2}s_t(n) \quad (55)\\ =& \sqrt{P_B\beta}\,\mathbf{h}_B s_B(n) + \sqrt{P_E\beta_2}\,\mathbf{h}_E s_E(n) \\ &+ (\mathbf{h}_{c2} - \hat{\mathbf{h}}_{c2})s_t(n) + \mathbf{v}(n) \quad (56)\\ \approx& \sqrt{P_B\beta}\,\mathbf{h}_B s_B(n) + \sqrt{P_E\beta_2}\,\mathbf{h}_E s_E(n) + \mathbf{v}(n). \quad (57)\end{aligned}$$

Since sequences $\{s_B(n)\}$ and $\{s_E(n)\}$ are independent non-Gaussian, one can apply higher-order statistics-based approaches to estimate $\mathbf{h}_B$ and $\mathbf{h}_E$ [28]–[31]. We first whiten $\check{\mathbf{y}}_2(n)$, and then use the RobustICA algorithm of [30] that uses kurtosis of "unmixed" whitened measurements. It provides estimates $\hat{\mathbf{h}}_B$ and $\hat{\mathbf{h}}_E$ of $\mathbf{h}_B$ and $\mathbf{h}_E$, respectively, up to a complex scaling factor, using $\check{\mathbf{y}}_2(n)$. Details follow.

Let

$$\mathbf{R}_{\check{y}_2} = \frac{1}{T}\sum_{n=1}^{T}\mathbb{E}\{\check{\mathbf{y}}_2(n)\check{\mathbf{y}}_2^H(n)\} \text{ and } \widehat{\mathbf{R}}_{\check{y}_2} = \frac{1}{T}\sum_{n=1}^{T}\check{\mathbf{y}}_2(n)\check{\mathbf{y}}_2^H(n), \quad (58)$$

where $\widehat{\mathbf{R}}_{\check{y}_2}$ is a consistent estimator of the correlation matrix $\mathbf{R}_{\check{y}_2}$ of $\{\check{\mathbf{y}}_2(n)\}$. Consider the EVD of $\widehat{\mathbf{R}}_{\check{y}_2}$ to obtain

$$\widehat{\mathbf{R}}_{\check{y}_2} = \hat{\mathbf{U}}\hat{\Sigma}\hat{\mathbf{U}}^H = [\hat{\mathbf{U}}_1\ \hat{\mathbf{U}}_2]\begin{bmatrix}\hat{\Sigma}_1 & \mathbf{0}\\ \mathbf{0} & \hat{\Sigma}_2\end{bmatrix}[\hat{\mathbf{U}}_1\ \hat{\mathbf{U}}_2]^H \quad (59)$$

where $\hat{\Sigma}$ is an $N_r \times N_r$ diagonal matrix with the eigenvalues of $\widehat{\mathbf{R}}_{\check{y}_2}$ arranged in decreasing order of magnitude, the columns of $\hat{\mathbf{U}}$ are the corresponding eigenvectors, and $\hat{\mathbf{U}}_1$ is $N_r \times 2$. With reference to (57), define a channel matrix $\mathbf{H}_d \in \mathbb{C}^{N_r \times 2}$ as

$$\mathbf{H}_d = [\sqrt{P_B \beta}\, \mathbf{h}_B \quad \sqrt{P_E \beta_2}\, \mathbf{h}_E]. \tag{60}$$

Then we can rewrite (57) as

$$\check{\mathbf{y}}_2(n) = \mathbf{H}_d \mathbf{s}(n) + \mathbf{v}(n), \quad \mathbf{s}(n) = [s_B(n)\ s_E(n)]^T. \tag{61}$$

Since the contamination sequences $s_B(n)$ and $s_E(n)$ are zero-mean, unit variance, mutually independent and i.i.d., we have the true correlation function

$$\begin{aligned} \mathbf{R}_{\check{y}_2} &= \mathbf{U}\Sigma\mathbf{U}^H = [\mathbf{U}_1\ \mathbf{U}_2] \begin{bmatrix} \Sigma_1 & \mathbf{0} \\ \mathbf{0} & \Sigma_2 \end{bmatrix} [\mathbf{U}_1\ \mathbf{U}_2]^H \\ &= \mathbf{H}_d \mathbf{H}_d^H + \sigma_v^2 \mathbf{I}_{N_r} \end{aligned} \tag{62}$$

where $\mathbf{U}$, Σ, etc. in (62) are the true counterparts of the estimated $\hat{\mathbf{U}}$, $\hat{\Sigma}$, etc. in (59).

The channels $\mathbf{h}_B$ and $\mathbf{h}_E$ lie in the subspace spanned by the columns of $\mathbf{U}_1$. Consider $\mathbf{x}(n) = \mathbf{U}_1^H \check{\mathbf{y}}_2(n) \in \mathbb{C}^2$. Then we have

$$\mathbf{x}(n) = \mathbf{U}_1^H (\mathbf{H}_d \mathbf{s}(n) + \mathbf{v}(n)) = \tilde{\mathbf{H}}_d \mathbf{s}(n) + \check{\mathbf{v}}(n) \tag{63}$$

where $\tilde{\mathbf{H}}_d \in \mathbb{C}^{2\times 2}$, $\tilde{\mathbf{H}}_d = \mathbf{U}_1^H \mathbf{H}_d$, and $\check{\mathbf{v}}(n) = \mathbf{U}_1^H \mathbf{v}(n) \in \mathbb{C}^2$. We have $\mathbb{E}\{\check{\mathbf{v}}(n)\check{\mathbf{v}}^H(n)\} = \sigma_v^2 \mathbf{I}_2$ since $\mathbf{U}_1^H \mathbf{U} = \mathbf{I}_2$. We will use the RobustICA algorithm of [30] to yield an estimate $\hat{\tilde{\mathbf{H}}}_d$ of $\tilde{\mathbf{H}}_d$ using $\mathbf{x}(n)$. One obtains, for some $\theta_i s$,

$$\hat{\tilde{\mathbf{H}}}_d \approx \tilde{\mathbf{H}}_d \mathcal{P} \Gamma_\theta, \quad \Gamma_\theta = \mathrm{diag}\{e^{j\theta_i},\ i = 1, 2\} \tag{64}$$

where $\mathcal{P}$ is a permutation matrix – the order of "extracted" sources, hence, the order of extracted columns of $\tilde{\mathbf{H}}_d$ cannot be determined by RobustICA (indeed, by any blind source separation method for instantaneous mixtures [31]), and one can only recover channels up to a constant of modulus one when using kurtosis and related criteria for unmixing. Thus, an estimate of $\mathbf{H}_d = \mathbf{U}_1 \tilde{\mathbf{H}}_d$ is given by

$$\hat{\mathbf{H}}_d = \hat{\mathbf{U}}_1 \hat{\tilde{\mathbf{H}}}_d = [\sqrt{P_B \beta}\, \hat{\mathbf{h}}_B \quad \sqrt{P_E \beta_2}\, \hat{\mathbf{h}}_E] \approx \mathbf{H}_d \mathcal{P} \Gamma_\theta\,. \tag{65}$$

Remark 1: Consider the (restricted) independent component analysis (ICA) problem

$$\mathbf{z}(n) = \mathbf{A}\mathbf{s}(n) \tag{66}$$

where $\mathbf{z}(n), \mathbf{s}(n) \in \mathbb{C}^p$, $\mathbf{A} \in \mathbb{C}^{p\times p}$, the sequence $\{\mathbf{s}(n)\}$ is zero-mean, i.i.d., non-Gaussian (finite alphabet), with independent components, and the objective is to recover $\mathbf{s}(n)$ and estimate $\mathbf{A}$. Such problems have been addressed in [28]–[31], among others. We will consider only square $\mathbf{A}$, hence the term restricted ICA; this is sufficient for our purposes. Ignoring noise $\check{\mathbf{v}}(n)$ in (63), we see that (66) corresponds to (63) with $p = 2$, $\mathbf{A} = \tilde{\mathbf{H}}_d$, and $\mathbf{z}(n) = \mathbf{x}(n)$. For some $\mathbf{w} \in \mathbb{C}^p$, let $e(n) = \mathbf{w}^H \mathbf{z}(n)$. In the approach of [30], $\mathbf{w}$ is picked to maximize $|\gamma_4|$, where the kurtosis (normalized 4th cumulant) γ_4 of $e(n)$ is given by

$$\gamma_4 = \frac{\mathbb{E}\{|e(n)|^4\} - 2\mathbb{E}\{|e(n)|^2\} - \mathbb{E}\{e^2(n)\}}{\left(\mathbb{E}\{|e(n)|^2\}\right)^2}. \tag{67}$$

When $|\gamma_4|$ is maximized for $\mathbf{w} = \bar{\mathbf{w}}$, one has $e(n) = \bar{\mathbf{w}}^H \mathbf{z}(n) = s_m(n)$ for some $1 \le m \le p$, where $s_m(n)$ is the mth component of $\mathbf{s}(n)$. Thus, one can obtain $s_m(n)$, and using $s_m(n)$, (66) and least-squares, estimate $A_{\ell m}$ for $1 \le \ell \le p$, where $A_{\ell m}$ is (ℓ, m)th element of $\mathbf{A}$. After estimating the mth column of $\mathbf{A}$, contribution of $s_m(n)$ to $\mathbf{z}(n)$ is subtracted (deflated) from $\mathbf{z}(n)$, and the entire process is repeated till all sources (components of $\mathbf{s}(n)$) are extracted, and $\mathbf{A}$ is estimated. Such a procedure for more general class of systems may also be found in [28]. The RobustICA algorithm of [30] provides an optimal step-size in the iterative maximization of $|\gamma_4|$ via a gradient-descent method, leading to fast convergence. □

Remark 2: Note that order of the extracted sources in an ICA problem is unknown [30], [31]. That is, in Remark 1, the index m in the recovered $s_m(n)$ could correspond to any of the existing sources in the mixture. Therefore, the channel estimated based on the extracted source signal is not "labeled," i.e., with reference to (57), we do not know if an estimated channel resulting from the application of the RobustICA algorithm of [30] (or any other unmixing approach), corresponds to that of Bob or of Eve. We need some additional information to resolve this ambiguity. If old, outdated estimates of Bob's channel (from earlier frames) are available to Alice, they can be used to distinguish between current estimates of Bob's and Eve's channel. If Bob and Eve use different symbol constellations for random sequences, it can help distinguish between the two. In this paper, we assume that either outdated estimates of Bob's channel are available, or the superimposed random sequence of Bob has some information embedded in it regarding user identification, and Alice can extract this from decoded data. □

A. Summary of Solution

Alice does not know if attack model (4) or (5) is in effect. Based on the discussion in Secs. III and IV, we have the following approach to pilot spoofing attack detection, and estimation of both legitimate receiver and eavesdropper channels.

(i) Apply RMT rank estimator (13) to (original) measurements $\mathbf{y}(n)$, to determine signal subspace rank $\widehat{d}_{RMT}$. If $\widehat{d}_{RMT} = 1$, estimate Bob's channel via the method of [12, Sec. IV].

(ii) If $\widehat{d}_{RMT} > 1$, there is a spoofing attack. Generate the projected sequence $\tilde{\mathbf{y}}(n)$ as specified in Sec. III-C. Apply the RMT rank estimator to $\{\tilde{\mathbf{y}}(n)\}$ to estimate signal subspace rank $\hat{\tilde{d}}$ of the projected measurements. If $\hat{\tilde{d}} = 1$, attack model (4) is in effect, and if $\hat{\tilde{d}} > 1$, attack model (5) applies.

(iii) If $\hat{\tilde{d}} = 1$, follow second-order statistics-based method discussed in Sec. III-C to obtain estimates of $\mathbf{h}_B$ and $\mathbf{h}_E$.

(iv) If $\hat{\tilde{d}} > 1$, follow kurtosis-based method discussed in Sec. IV to obtain estimates $\hat{\mathbf{h}}_B$ and $\hat{\mathbf{h}}_E$.

V. Countermeasure: Secure Beamforming

In this section, we discuss measures taken at Alice for her transmission to Bob, taking into account presence/absence of Eve, or simply ignoring Eve. These measures are based on known results, but their applicability depends upon the spoofing detection results and consequent estimation of the channels of Bob and Eve, as appropriate.

Let $\{s_A(n)\}$, $\mathbb{E}\{|s_A(n)|^2\} = 1$, denote the scalar information sequence of Alice intended for Bob, and let $\mathbf{w} \in \mathbb{C}^{N_r}$ denote the unit norm beamforming vector of Alice. Then Alice transmits $\sqrt{P_A}\mathbf{w}\, s_A(n)$ where P_A is the transmit power. The received signals at Bob and Eve are given, respectively, by

$$y_B(n) = \sqrt{P_A}\mathbf{h}_B^\top \mathbf{w}\, s_A(n) + v_B(n) \tag{68}$$

$$y_{AE}(n) = \sqrt{P_A}\mathbf{h}_E^\top \mathbf{w}\, s_A(n) + v_E(n), \tag{69}$$

where we have used channel reciprocity, $v_E(n) \sim \mathcal{N}_c(0, \sigma_E^2)$ and $v_B(n) \sim \mathcal{N}_c(0, \sigma_B^2)$ are additive white Gaussian noise at Eve's and Bob's receivers.

A. MF Beamforming

For matched filter (MF) reception at Bob, Alice should pick $\mathbf{w}$ as $\mathbf{h}_B^*/\|\mathbf{h}_B\|$ if $\mathbf{h}_B$ is known [32], [33], but instead uses the estimated channel to pick the optimum beamformer

$$\mathbf{w}_* = \hat{\mathbf{h}}_B^*/\|\hat{\mathbf{h}}_B\|. \tag{70}$$

The choice $\mathbf{w} = \mathbf{h}_B^*/\|\mathbf{h}_B\|$ maximizes the SNR at Bob since $|\mathbf{h}_B^\top \mathbf{w}| \le \|\mathbf{h}_B\|\,\|\mathbf{w}\|$ with equality iff $\mathbf{w} = c\mathbf{h}_B^*$ for some constant c. We will refer to the solution (70) as MF beamforming. Generally, we would use MF beamforming when there is no Eve (or Eve is not detected) and we use the estimate of Bob's channel as obtained in Sec. III-D1.

B. Blind Beamforming

In blind beamforming, we do not depend upon the decision of the attack detector, and use MF beamformer based on blind estimate of Bob's channel, which is obtained as described in Sec. III-D1. Since the blind channel estimate is obtained from projected data, it is not influenced by the presence or absence of Eve's signal.

C. Secure Beamforming

We now discuss the case where Eve has been detected. Using beamformer $\mathbf{w}$ at Alice, the signal-to-noise ratios (SNRs) SNR_B and SNR_E at Bob and Eve, respectively, are

$$\mathrm{SNR}_B = P_A|\mathbf{h}_B^\top \mathbf{w}|^2/\sigma_B^2, \quad \mathrm{SNR}_E = P_A|\mathbf{h}_E^\top \mathbf{w}|^2/\sigma_E^2. \tag{71}$$

If a Gaussian codebook is used for $\{s_A(n)\}$, the achievable rates at Bob and Eve, respectively, are $R_B = \log_2(1 + \mathrm{SNR}_B)$ and $R_E = \log_2(1 + \mathrm{SNR}_E)$, and the secrecy rate at Bob is [1]

$$R_{B,sec} = \max(R_B - R_E, 0). \tag{72}$$

In the presence of Eve with channel $\mathbf{h}_E$, the beamformer $\mathbf{w}$ may be picked to maximize $R_{B,sec}$. By [1, Theorem 2], the optimal beamformer $\mathbf{w}_*$ is given by the (unit-norm) generalized eigenvector corresponding to the largest generalized eigenvalue of the matrix pair

$$\left(\mathbf{I}_{N_r} + \mathbf{h}_B^*\mathbf{h}_B^\top/\sigma_B^2,\, \mathbf{I}_{N_r} + \mathbf{h}_E^*\mathbf{h}_E^\top/\sigma_E^2\right). \tag{73}$$

Under high SNR, the above solution approaches the solution that satisfies $\mathbf{h}_E^\top \mathbf{w} = 0$ [1, Cor. 1] (see the discussion in the second paragraph after the statement of Cor. 1 in [1], and also its proof in [1, Sec. VI.B]). If $\mathbf{h}_E^\top \mathbf{w} = 0$, then $R_E = 0$ and maximizing $R_{B,sec}$ is equivalent to maximization of rate R_B, hence of $|\mathbf{h}_B^\top \mathbf{w}|$. This leads to the optimization problem

$$\max_{\mathbf{w}} \ |\mathbf{h}_B^\top \mathbf{w}| \quad \text{subject to } \ \mathbf{h}_E^\top \mathbf{w} = 0,\ \|\mathbf{w}\| = 1. \tag{74}$$

The constraint $\mathbf{h}_E^\top \mathbf{w} = 0$ implies that $\mathbf{w}$ lies in a subspace orthogonal to $\mathbf{h}_E^*$, i.e., for some $\mathbf{w}_0$, with $\mathcal{P}^\perp_{\mathbf{h}_E^*}$ denoting projection orthogonal to $\mathbf{h}_E^*$,

$$\mathbf{w} = \mathcal{P}^\perp_{\mathbf{h}_E^*}\mathbf{w}_0 = \left(\mathbf{I}_{N_r} - \mathbf{h}_E^*\mathbf{h}_E^\top/\|\mathbf{h}_E\|^2\right)\mathbf{w}_0. \tag{75}$$

With $\tilde{\mathbf{h}}_B := (\mathcal{P}^\perp_{\mathbf{h}_E^*})^\top \mathbf{h}_B$, $|\mathbf{h}_B^\top \mathbf{w}| = |\tilde{\mathbf{h}}_B^\top \mathbf{w}_0|$ is maximized w.r.t. $\mathbf{w}_0$ by an MF solution $\mathbf{w}_{0*} = c\,\tilde{\mathbf{h}}_B^*$ for some nonzero constant c, as in Sec. V-A. Since $\mathcal{P}^\perp_{\mathbf{h}_E^*}$ is a projection operator satisfying $\mathcal{P}^\perp_{\mathbf{h}_E^*}(\mathcal{P}^\perp_{\mathbf{h}_E^*})^H = \mathcal{P}^\perp_{\mathbf{h}_E^*}$, in terms of $\mathbf{w}$, we have $\mathbf{w} = \mathcal{P}^\perp_{\mathbf{h}_E^*}\mathbf{w}_{0*} = c\mathcal{P}^\perp_{\mathbf{h}_E^*}\mathbf{h}_B^*$, which after normalization to unit norm, leads to the optimal solution

$$\mathbf{w}_* = \frac{\left(\mathbf{I}_{N_r} - \mathbf{h}_E^*\mathbf{h}_E^\top/\|\mathbf{h}_E\|^2\right)\mathbf{h}_B^*}{\|\left(\mathbf{I}_{N_r} - \mathbf{h}_E^*\mathbf{h}_E^\top/\|\mathbf{h}_E\|^2\right)\mathbf{h}_B^*\|}. \tag{76}$$

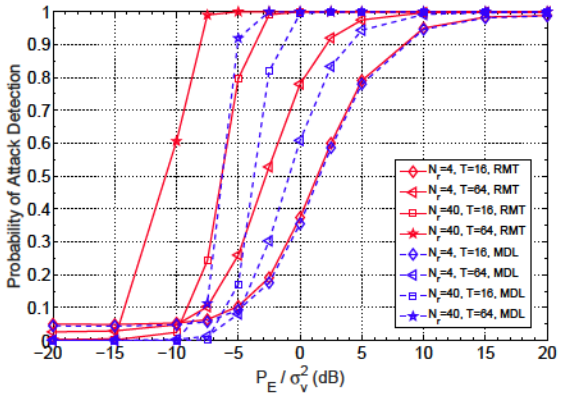
The optimization problem (74) has been investigated in [34] in a different context, with solution (76). In practice, we replace $\mathbf{h}_B$ and $\mathbf{h}_E$ with their estimates. We will refer to the optimal solution (73) and suboptimal solution (76), both as secure beamforming.

VI. Simulation Examples

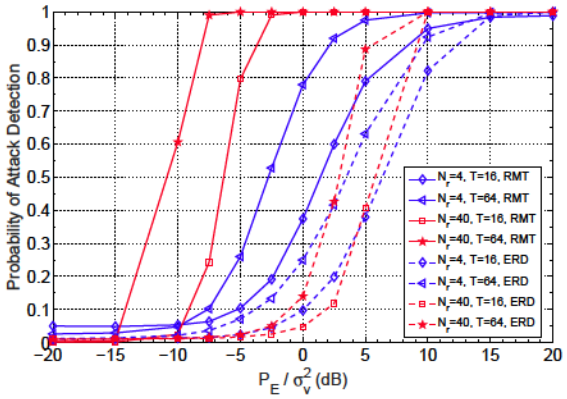
We consider Rayleigh flat-fading channels with path losses $d_B = d_E = 1$, noise power σ_v^2, training power budget P_B at Bob is such that $P_B/\sigma_v^2 = 10$dB, training power budget P_E at Eve is such that P_E/σ_v^2 varies from -20dB through 20dB, and fractional allocation β of training power at Bob to random sequence $s_B(n)$ is 0.4, and $\beta_2 = 0.4$ when Eve chooses to add $s_E(n)$ under attack model (5). Bob and Eve have single antennas while Alice has $N_r = 4$ or 40 antennas. The training sequence is selected as periodic extension of a (binary) Hadamard sequence of length $P = 2^4 = 16$, and the random sequences $\{s_B(n)\}$ and $\{s_E(n)\}$ were i.i.d. QPSK symbols. Alice knows that $\{s_B(n)\}$ is QPSK, but does not know the alphabet size of $\{s_E(n)\}$. All shown simulation results were averaged over 5000 runs.

A. Attack Model (4)

Here we consider the case where Eve does not transmit any random sequence, just the pilot. Figs. 1a-1b show our detection probability P_d results averaged over 5000 runs under pilot contamination attack for various parameter choices, when $P_B/\sigma_v^2 = 10dB$. Fig. 1a compares RMT and MDL

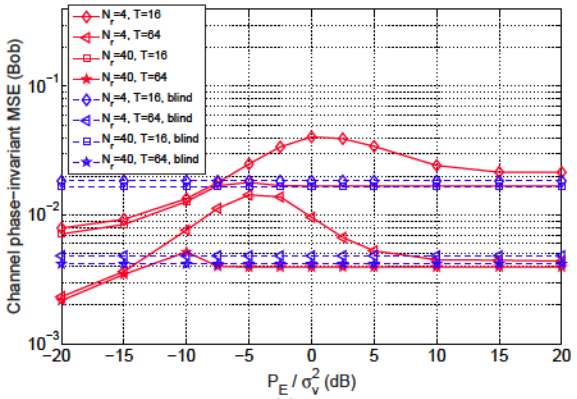


(a)

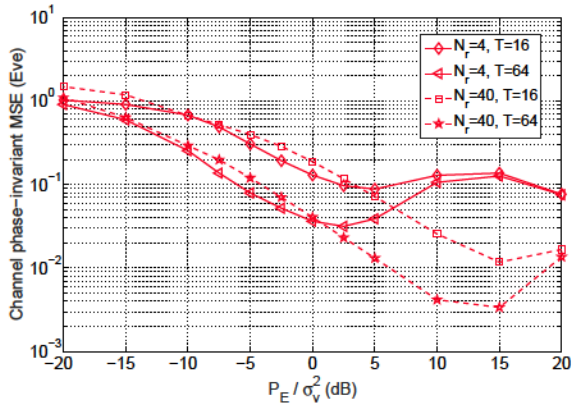


(b)

Fig. 1: *Attack model (4). (a) Probability of attack detection using MDL or RMT estimators, as a function of Eve's power P_E relative to noise power σ_v^2, when Bob's power is fixed at $P_B/\sigma_v^2 = 10$dB. $\beta = 0.4$.*
(b) Probability of attack detection using RMT estimator or energy-ratio detector (ERD) of [10], as a function of Eve's power P_E relative to noise power σ_v^2, when Bob's power is fixed at $P_B/\sigma_v^2 = 10$dB. $\beta = 0.4$.



(a)



(b)

Fig. 2: *Attack model (4). Phase-insensitive channel MSEs as a function of Eve's power P_E relative to noise power σ_v^2, when Bob's power is fixed at $P_B/\sigma_v^2 = 10$dB. $\beta = 0.4$. (a) MSE (77) for Bob's channel. The method of Sec. III-B was used for curves labeled "blind" and indicated via dashed curves. For solid curves, we use the method of Sec. III-C if the RMT detector indicates the presence of Eve's attack, otherwise use the method of Sec. III-A .*
(b) MSE (78) for Eve's channel. We use the method of Sec. III-C if the RMT detector indicates the presence of Eve's attack, else Eve's channel is not estimated.

approaches. We picked $P_{fa} = 0.001$ for the RMT method. The MDL method needs no threshold. It is seen from Fig. 1a that the RMT method significantly outperforms the MDL detector, in that the RMT method is able to detect pilot spoofing attack for much smaller values of Eve's power P_E compared to MDL. For the cases shown in Fig. 1a, the empirical P_{fa} for the MDL method was less than 0.0008 for $(N_r, T) = (40, 16)$, (4,64) and (40,64), and it was 0.048 for $(N_r, T) = (4, 16)$, whereas for the RMT method, the empirical probability of false alarm was 0.049, 0.026, 0.002 and 0.002, for $(N_r, T) = (4, 16)$, (4,64), (40,16) and (40,64), respectively. Fig. 1b compares RMT-based detector with the energy-ratio detector (ERD) of [10], the latter designed for $P_{fa} = 0.05$, which additionally requires feedback from Alice to Bob. It is seen from Fig. 1b that the RMT method significantly outperforms ERD. Also, comparing Figs. 1a and 1b we see that the MDL detector also outperforms ERD. For the cases shown in Fig. 1b, the empirical P_{fa} for the RMT method was exactly as for Fig. 1a, whereas for ERD, it was 0.050, 0.044, 0.048 and 0.058, for $(N_r, T) = (4, 16)$, (4,64), (40,16) and (40,64), respectively.

Given the channel $\mathbf{h}_B$ and its rotated version $e^{j\theta}\mathbf{h}_B$ for some θ, the solution to secure beamforming via (73) or (76), and the solution to MF or blind beamforming via (70), none depends upon the phase rotation θ. Similar comments apply to $\mathbf{h}_E$ and its rotated version $e^{j\theta}\mathbf{h}_E$ when secure beamforming is considered. Therefore, we will use a phase-insensitive mean-square error (MSE) measure to evaluate channel estimation errors in estimating $\mathbf{h}_B$ and $\mathbf{h}_E$. Such a measure has been used in [35] in a different context. Note also that (70) and (76) do not depend upon the norms of $\mathbf{h}_B$ and $\mathbf{h}_E$. If $\hat{\mathbf{h}}_B$ is an estimate of $\mathbf{h}_B$, both normalized to unit norm, phase-insensitive MSE in estimation of $\mathbf{h}_B$ is given by [35]

$$\text{CMSE}_B = \min_{\theta\in[0,2\pi]} \|\mathbf{h}_B - e^{j\theta}\hat{\mathbf{h}}_B\|^2 = 2 - 2|\mathbf{h}_B^H\hat{\mathbf{h}}_B|. \quad (77)$$

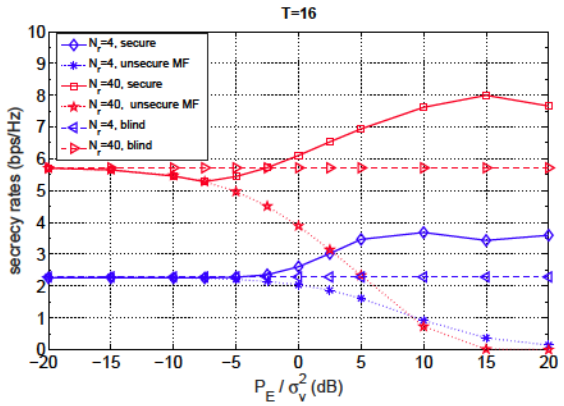
Similarly, if $\hat{\mathbf{h}}_E$ is an estimate of $\mathbf{h}_E$, both normalized to unit norm, phase-insensitive MSE in estimation of $\mathbf{h}_E$ is given by

$$\text{CMSE}_E = \min_{\theta\in[0,2\pi]} \|\mathbf{h}_E - e^{j\theta}\hat{\mathbf{h}}_E\|^2 = 2 - 2|\mathbf{h}_E^H\hat{\mathbf{h}}_E|. \quad (78)$$

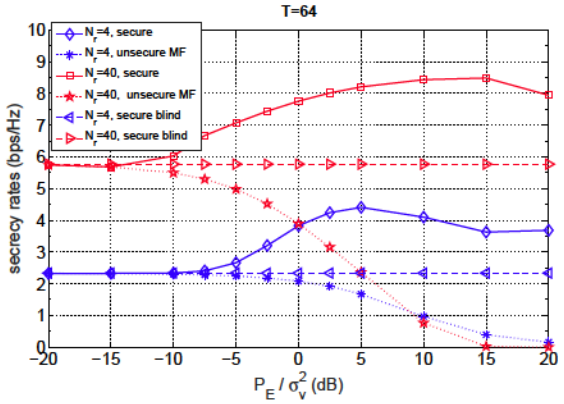
Figs. 2a and 2b show phase-insensitive MSE in Bob's and Eve's channel estimation, respectively. Bob's channel can be estimated two different ways: blind estimation as discussed in Sec. III-B (labeled "blind" and indicated via dashed curves in Fig. 2a), or estimation as in Sec. III-C only if the RMT method indicates the presence of Eve's attack, otherwise use the method of Sec. III-A (shown via solid curves in Fig. 2a). Since the blind approach is applied after projection of data orthogonal to training, independent of the result of the attack detector, and since the projected data is devoid of Eve's signal, the MSE in Bob's blind channel estimation is invariant to Eve's power. This fact is reflected in Fig. 2a. On the other hand, when Bob's channel is estimated based on the result of the attack detector (methods of Sec. III-A or Sec. III-C), the channel estimation performance depends upon Eve's power. If the RMT method indicates absence of the attack, we use the method of Sec. III-A, where Eve's signal at Alice is not modeled, and therefore, it contributes to increased noise, resulting in increasing MSE with increasing Eve's power. However, as soon as the attack is detected, we switch to the method of Sec. III-C, where Eve's signal at Alice is explicitly modeled, resulting in improved channel MSE in Fig. 2a with increasing P_E. As T increases, we have improved attack detection (see Fig. 1a), which results in lower channel MSE in Fig. 2a with increasing T.

Eve's channel is estimated only if the RMT detector indicates the presence of a spoofing attack. The phase-insensitive MSE CMSE_E in Eve's channel estimation is shown in Fig. 2b. At first, it decreases with increasing P_E, before leveling off, or increasing a little. Since, in Sec. III-C, $\mathbf{h}_E$ is estimated after canceling the contribution of Bob's signal from $\mathbf{y}(n)$, Bob's channel needs to be estimated with accurate phase rotation and magnitude, to yield "effective" cancellation. It turns out that, at higher values of P_E, such errors increase, which, in turn, cause leveling off or increase in CMSE_E shown in Fig. 2b. This discrepancy, however, lessens with increasing N_r.

Based on the channel estimates, Alice designs secure beamformers for downlink transmission to Bob, as discussed in Sec. V. Bob's secrecy rate results are shown in Figs. 3a-3b for $T = 16$ and $T = 64$, respectively. For comparison, we also show the results for the case where Alice is ignorant of Eve's attack, and estimates Bob's channel as described in Sec. III-A, and then designs MF beamformer (70). These results are labeled "unsecure MF" in Figs. 3a-3b. It is seen that as Eve's spoofing power P_E increases, the secrecy rates decrease to zero for all values of T and N_r, since the estimated Bob's channel at Alice is now dominated by Eve's channel. There is significant information leakage to Eve. Now consider the case where Alice is aware that there could be Eve's pilot spoofing attack, but instead of trying to detect it, simply uses blind estimation of Bob's channel (Sec. III-B), and then designs MF beamformer (70). These results are labeled "blind" in Figs. 3a-3b. For this case, it is seen from Figs. 3a-3b that the secrecy rates are invariant to Eve's spoofing power P_E, resulting in secure transmission to Bob. The curves labeled "secure" refer to beamforming based on (73) when Eve is detected and on (70) when Eve is not detected, and we use the methods of Sec. III-A or Sec. III-C for channel estimation, depending upon if the attack is not detected or detected. In our simulations, we did not find any discernible difference between secure beamformers based on (73) or (76). As seen in Figs. 3a-3b, when Eve increases her transmit power P_E, she reveals more of herself (i.e., leads to better channel estimation performance at Alice, see Fig. 2b), resulting in better secrecy rate for Bob with increasing P_E. We achieve better nulling by the beamformer of Alice along Eve's direction.



(a)

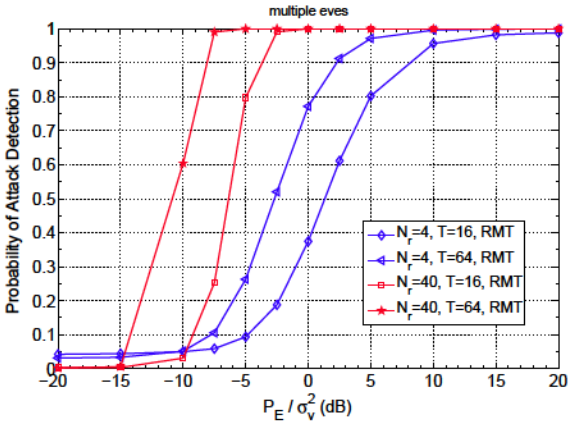


(b)

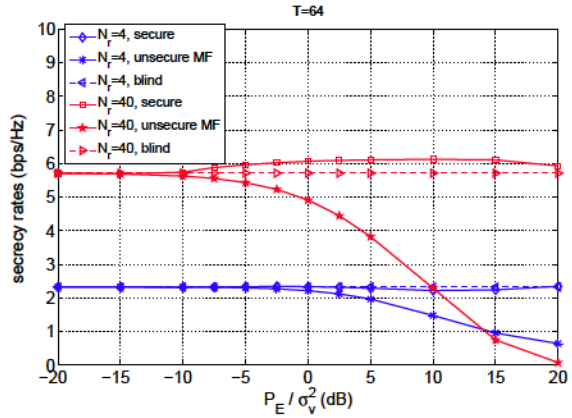
Fig. 3: *Attack model (4). Secrecy rate (bps/Hz) at Bob using the beamformers based on (70) or (73), as a function of Eve's power P_E. All parameters as for Fig. 1a. The label "secure" refers to beamforming based on (73) when Eve is detected, and on (70) when Eve is not detected; "unsecure MF" means one uses (70) with Eve ignored in channel estimation, i.e., Bob's channel is estimated using the method of Sec. III-A; "blind" means one uses (70) with Bob's channel estimated using the method of Sec. III-B. $P_A = 1$, $\sigma_B^2 = \sigma_E^2 = 0.1$. (a) $T = 16$, (b) $T = 64$.*

Multiple Eves: We also considered the case of 3 Eves under attack model (4) where (2) is modified as $\mathbf{y}(n) = \left(\sqrt{P_B}\,\mathbf{h}_B + \sum_{i=1}^{3}\sqrt{P_{E_i}}\,\mathbf{h}_{E_i}\right) s_t(n) + \mathbf{v}(n)$, and $\mathbf{h}_{E_i}$s are mutually independent. We take $P_{E_i} = P_E/3\ \forall i$. This model is tantamount to having one virtual Eve with $\sqrt{P_E}\,\mathbf{h}_E = \sum_{i=1}^{3}\sqrt{P_{E_i}}\,\mathbf{h}_{E_i}$. Our spoofing detection results are shown in Fig. 4a, and they are similar to the single Eve results of Fig. 1a. After obtaining the channel estimates $\hat{\mathbf{h}}_B$ and $\hat{\mathbf{h}}_E$ (we can not get $\hat{\mathbf{h}}_{E_i}$s), we design secure beamformers as for the results in Fig. 3a ($T = 64$), and evaluate secrecy rates of Bob (shown in Fig. 4b) w.r.t. one of the Eves (in downlink each

Eve receives its own Alice-to-Bob signal). Since now Alice does not know a single Eve's channel with any accuracy, the secrecy rate does not improve with increasing P_E (compare with Fig. 3a), but it does not deteriorate either when using blind or secure beamformers.



(a)



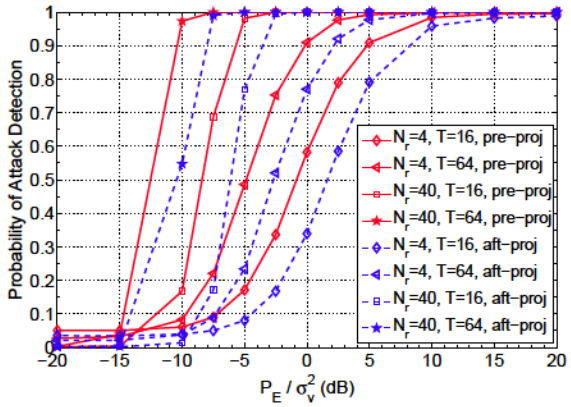
(b)

Fig. 4: *Multiple Eves in attack model (4). (a) Probability of attack detection using RMT estimators, as a function of three Eves' total power P_E relative to noise power σ_v^2, when Bob's power is fixed at $P_B/\sigma_v^2 = 10dB$. $\beta = 0.4$.*
(b) Secrecy rate (bps/Hz) at Bob using the beamformers based on (70) or (73), as a function of three Eves' total power P_E. The labels are as in Fig. 3a.

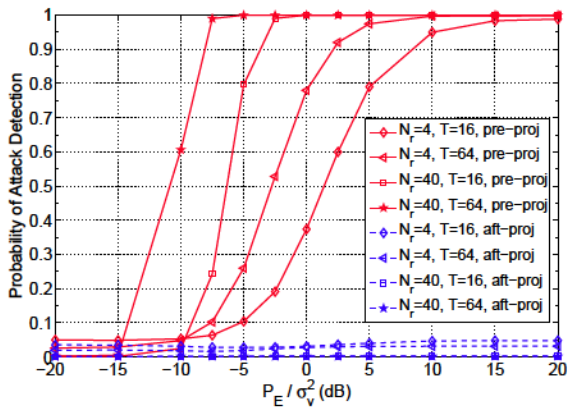
B. Attack Model (5)

Here we consider the case where Eve also transmits a random sequence, in addition to the pilot. Alice follows the algorithm detailed in Sec. IV-A. We use $\beta = \beta_2 = 0.4$, i.e., Eve behaves just like Bob. Fig. 5 shows our spoofing detection results using the RMT-based detector.. The label "pre-proj" means that the detector operates on original $\mathbf{y}(n)$, and label "aft-proj" means that the detector operates on projected $\tilde{\mathbf{y}}(n)$, given by (22) or (51). Fig. 5a shows the results for data generated under attack model (5) whereas Fig. 5b shows the results for data generated under attack model (4). Our detection approach (step (ii) in Sec. IV-A) clearly distinguishes between models (4) and (5). Note that Alice has no prior knowledge as to which attack model is true.

Figs. 6 and 7 show the channel estimation and secrecy rate results, respectively, for data generated under attack model (5), and they correspond to Figs. 2 and 3, respectively, for attack model (4). Under attack model (5), even if Eve's presence is correctly detected at low P_E/σ_v^2 values, since the channel estimation performance depends upon the relative strength of Bob's and Eve's random signals, and on T (in general, higher-order statistics-based approaches need larger data samples), there are larger errors in Eve's channel estimation at lower values of Eve's power P_E/σ_v^2 at Alice. Compare Figs. 2b and 6b for P_E/σ_v^2 values between -7.5dB to 5dB for $T = 16$, to notice this phenomenon. Poor estimation of Eve's channel seen in Fig. 6b is reflected in Fig. 7a, in decrease in Bob's secrecy rate for these values of P_E/σ_v^2 and $T = 16$, since poorer Eve's channel estimates lead to poorer nulling by the beamformer of Alice along Eve's direction.



(a)

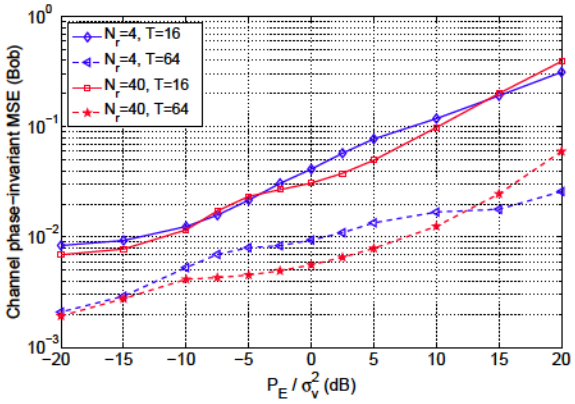


(b)

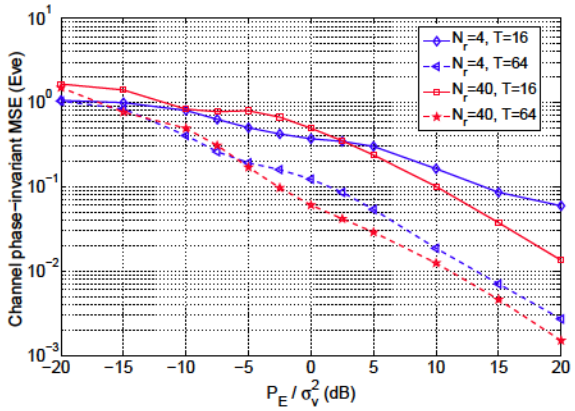
Fig. 5: *Probability of attack detection using RMT rank estimator, as a function of Eve's power P_E relative to noise power σ_v^2, when Bob's power is fixed at $P_B/\sigma_v^2 = 10dB$. (a) Attack model (5), $\beta = \beta_2 = 0.4$, (b) Attack model (4), $\beta = 0.4$, $\beta_2 = 0$.*

VII. Conclusion

A novel approach to detection of pilot spoofing attack in a three-node TDD system was recently presented in [12] where attack mitigation was not addressed. In this paper we augment the approach of [12] with estimation of the channels of Bob and Eve, followed by secure beamforming, to mitigate the



(a)



(b)

Fig. 6: *Attack model (5). Phase-insensitive channel MSEs as a function of Eve's power P_E relative to noise power σ_v^2, when Bob's power is fixed at $P_B/\sigma_v^2 = 10$dB. The method of Sec. IV was used. $\beta = \beta_2 = 0.4$. (a) MSE (77) for Bob's channel, (b) MSE (78) for Eve's channel.*

effects of pilot spoofing. Two spoofing attack models were considered: (i) the spoofer Eve transmits only the pilot signal, (ii) Eve also adds a random sequence to its pilot, mimicking Bob. We also employed a random matrix theory (RMT) based source enumeration approach, instead of the MDL method used in [12], for spoofing detection. The proposed approaches were illustrated by numerical examples.

The proposed approach is confined to a single Bob and single Eve. It is desirable to extend the results to the case of multiple Bobs and multiple Eves. The problem of pilot spoofing detection (but not of countermeasures) for multiple Bobs and Eves has been considered in [21], where each Bob uses an independent self-contamination signal, in addition to its assigned orthogonal pilot sequence. Although these self-contamination signals act as interference to other Bobs, it is shown in [21] that by using iterative methods for multi-user channel estimation (first estimate the channels based solely on orthogonal training sequences with self-contamination signals acting as noise/interference, then use a linear MMSE equalizer based on the estimated channels to estimate and quantize the self-contamination signals, and repeat with training plus quantized self-contamination signal acting as pseudo-training for each individual Bob), one can obtain successful CSI estimates at Alice for those Bobs that are not spoofed. In [21] an approach is presented for detection of pilot spoofing, identification of which pilot has been spoofed, and then estimation of the channels of unspoofed Bobs via an iterative approach. However, [21] does not address the issue of estimation of a particular Bob's channel when that Bob is under a pilot spoofing attack. This is an open problem.

References

[1] A. Khisti and G. Wornell, "Secure transmission with multiple antennas - I: The MISOME wiretap channel," *IEEE Trans. Inf. Theory,* vol. 56, pp. 3088-3104, July 2010.

[2] Y.-W. Hong, P.-C. Lan and C-C. Jay Kuo, "Enhancing physical-layer secrecy in multiantenna wireless systems: an overview of signal processing approaches," *IEEE Signal Proc. Mag.*, vol. 30, Issue 5, pp. 29-40, Sept. 2013.

[3] X. Zhou, B. Maham and A. Hjorungnes, "Pilot contamination for active eavesdropping," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 903-907, March 2012.

[4] D. Kapetanovic, G. Zheng, K-K. Wong and B. Ottersten, "Detection of pilot contamination attack using random training and massive MIMO," in *Proc. 2013 IEEE 24th Intern. Symp. Personal, Indoor, Mobile Radio Commun. (PIMRC)*, London, UK, Sept. 8-11, 2013, pp. 13-18.

[5] D. Kapetanovic, A. Al-Nahari, A. Stojanovic and F. Rusek, "Detection of active eavesdroppers in massive MIMO," in *Proc. 2014 IEEE 24th Intern. Symp. Personal, Indoor, Mobile Radio Commun. (PIMRC)*, Washington, DC, Sept. 2014, pp. 585-589.

[6] 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, June 2015.

[7] R. Miller and W. Trappe, "On the vulnerabilities of CSI in MIMO wireless communication systems," *IEEE Trans. Mobile Computing*, vol. 8, pp. 1386-1398, Aug. 2012.

[8] X. Chen, J. Chen, H. Zhang, Y. Zhang and C. Yuen, "On secrecy performance of multiantenna-jammer-aided secure communications with imperfect CSI," *IEEE Trans. Veh. Tech.*, vol. 65, no. 10, pp. 8014-8024, Oct. 2016.

[9] T.T. Do, E. Björnson, E.J. Larsson and S.M. Razavizadeh, "Jamming-resistant receivers for the massive MIMO uplink," *IEEE Trans. Inf. Forensics Security*, vol. 13, pp. 210-223, Jan. 2018.

[10] Q. Xiong, Y-C. Liang, K.H. Li and Y. Gong, "An energy-ratio-based approach for detecting pilot spoofing attack in multiple-antenna systems," *IEEE Trans. Inf. Forensics Security*, vol. 10, pp. 932-940, May 2015.

[11] S. Im, H. Jeon, J. Choi and J. Ha, "Secret key agreement with large antenna arrays under the pilot contamination attack," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 6579-6594, Dec. 2015.

[12] J.K. Tugnait, "Self-contamination for detection of pilot contamination attack in multiple antenna systems ," *IEEE Wireless Commun. Lett.*, vol. 4, No. 5, pp. 525-528, Oct. 2015.

[13] J.-M. Kang, C. In and H.-M. Kim, "Detection of pilot contamination attack for multi-antenna based secrecy systems," in *Proc. IEEE Veh. Tech. Conf. (VTC'15- Spring)*, Glasgow, Scotland, May 2015, pp. 1-5.

[14] Q. Xiong, Y-C. Liang, K.H. Li and Y. Gong, "Secure transmission against pilot spoofing attack: A two-way training-based scheme," *IEEE Trans. Inf. Forensics Security*, vol. 11, pp. 1017-1026, May 2016.

[15] J. Xie, Y-C. Liang,J. Fang and X. Kang, "Two-stage uplink training for pilot spoofing attack detection and secure transmission," in *Proc. 2017 IEEE Intern. Commun. Conf. (ICC)*, Paris, France, May 2017, pp. 1-6.

[16] X. Tian, M. Li, G. Ti, W. Liu and Q. Liu, "Random training-aided pilot spoofing detection," in *Proc. IEEE Wireless Commun. Signal Process. (WCSP) Conf.*, Yangzhou, China, Oct. 2016, pp. 1-5.

[17] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust. Speech, Signal Process.*, vol. 33, no. 2, pp. 387-392, April 1985.

[18] F. Haddadi, M. Malek-Mohammadi, M.M. Nayebi and M.R. Aref, "Statistical performance analysis of MDL source enumeration in array processing," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 452-457, Jan. 2010.

[19] B. Nadler, "Nonparametric detection of signals by information theoretic criteria: Performance analysis and an improved estimator," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2746-2756, May 2010.

[20] J.K. Tugnait, "Detection of pilot contamination attack in TDD/SDMA systems," in *Proc. 2016 IEEE Intern. Conf. Acoust., Speech Signal Process. (ICASSP 2016)*, Shanghai, China, March 20-25, 2016, pp. 3576-3580.

[21] J.K. Tugnait, "Detection and identification of spoofed pilots in TDD/SDMA systems," *IEEE Wireless Commun. Lett.*, vol. 6, no. 4, pp. 550-553, Aug. 2017.

[22] F. Bai, P. Ren, Q. Du and L. Sun, "A hybrid channel estimation strategy against pilot spoofing attack in MISO system," in *Proc. 2016 IEEE 24th Intern. Symp. Personal, Indoor, Mobile Radio Commun. (PIMRC)*, Valencia, Spain, Sept. 2016, pp. 1-6.

[23] J.K. Tugnait, "On mitigation of pilot spoofing attack," in *Proc. 2016 IEEE Intern. Conf. Acoust., Speech Signal Process. (ICASSP 2017)*, New Orleans, Louisiana, March 5-9, 2017, pp. 2097-2101.

[24] S. Kritchman and B. Nadler, "Determining the number of components in a factor model from limited noisy data," *Chem. Inst. Lab. Syst.*, vol. 94, pp. 19-32, 2008.

[25] S. Kritchman and B. Nadler, "Non-parametric detection of the number of signals: Hypothesis testing and random matrix theory," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3930-3941, Oct. 2009.

[26] N. El Karoui, "A rate of convergence result for the largest eigenvalue of complex white Wishart matrices," *Ann. Prob.*, vol. 36, no. 6, pp. 2077-2117, 2006.

[27] H.Q. Ngo and E.G. Larsson, "EVD-based channel estimation in multi-cell multiuser MIMO systems with very large antenna arrays," in *Proc. 2012 IEEE Intern. Conf. Acous., Speech & Signal Process. (ICASSP 2012)*, Kyoto, Japan, March 2012, pp. 3249-3252.

[28] J.K. Tugnait, "Identification and deconvolution of multichannel linear non-Gaussian processes using higher-order statistics and inverse filter criteria," *IEEE Trans. Signal Process.*, vol. 45, pp. 658-672, March 1997.

[29] J.F. Cardoso,"Blind signal separation: Statistical principles," *Proc. IEEE*, vol. 86, pp. 2009-2025, Oct. 1998.

[30] V. Zarzoso and P. Comon, "Robust independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size," *IEEE Trans. Neural Netw.*, vol. 21, pp. 248-261, Feb. 2010.

[31] P. Comon and C. Jutten, *Handbook of Blind Source Separation*. New York: Academic, 2010

[32] L. Lu, G.Y. Li, A.L. Swindlehurst, A. Ashikhmin and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742-758, Oct. 2014.

[33] T. Lo, "Maximal ratio transmission," *IEEE Trans. Commun.*, vol. 47, no. 10, pp. 1458-1461, Oct. 1999.

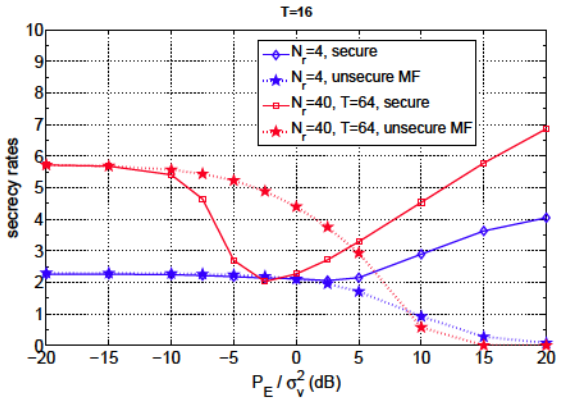
[34] A. El Shafie, D. Niyato and N. Al-Dhahir, "Security of rechargeable energy-harvesting transmitters in wireless networks," *IEEE Wireless Commun. Lett.*, vol. 5, no. 4, pp. 384-387, Aug. 2016.

[35] D.J. Love and R.W. Heath, "Equal gain transmission in multiple-input multiple-output wireless systems," *IEEE Trans. Commun.*, vol. 51, no. 7, pp. 1102-1110, July 2003.

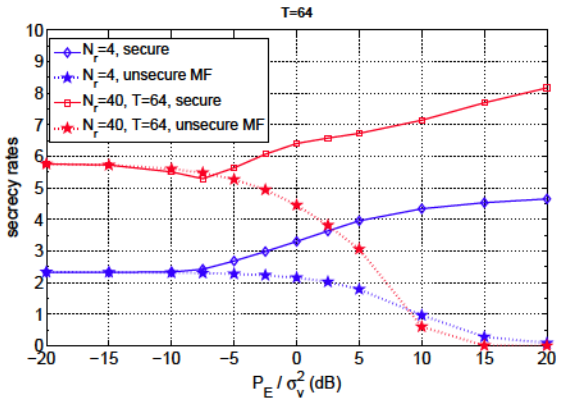
Jitendra K. Tugnait (M'79–SM'93–F'94) received the B.Sc.(Hons.) degree in electronics and electrical communication engineering from the Punjab Engineering College, Chandigarh, India in 1971, the M.S. and the E.E. degrees from Syracuse University, Syracuse, NY and the Ph.D. degree from the University of Illinois, Urbana-Champaign in 1973, 1974, and 1978, respectively, all in electrical engineering.

From 1978 to 1982 he was an Assistant Professor of Electrical and Computer Engineering at the University of Iowa, Iowa City, IA. He was with the Long Range Research Division of the Exxon Production Research Company, Houston, TX, from June 1982 to Sept. 1989. He joined the Department of Electrical & Computer Engineering, Auburn University, Auburn, AL, in September 1989 as a Professor. He currently holds the title of James B. Davis Professor. His current research interests are in statistical signal processing, wireless communications, and multiple target tracking.

Dr. Tugnait is a past Associate Editor of the IEEE Transactions on Automatic Control, the IEEE Transactions on Signal Processing, IEEE Signal Processing Letters, and the IEEE Transactions on Wireless Communications, past Senior Area Editor of the IEEE Transactions on Signal Processing, and past Senior Editor of IEEE Wireless Communications Letters.



(a)



(b)

Fig. 7: *Attack model (5). Secrecy rate (bps/Hz) at Bob using the beamformers based on (70) or (73), as a function of Eve's power* P_E. $P_A = 1$, $\sigma_B^2 = \sigma_E^2 = 0.1$ *(a)* $T = 16$, *(b)* $T = 64$.