Detection and Mitigation of Pilot Spoofing Attack

Jitendra K. Tugnait
Dept. of Electrical & Computer Eng.
Auburn University, Auburn, AL 36849, USA

***Abstract*—In a time-division duplex (TDD) multiple antenna system, the channel state information (CSI) can be estimated using reverse training. A pilot contamination (spoofing) attack occurs when during the training phase, an adversary (spoofer) also sends identical training (pilot) signal as that of the legitimate receiver. This contaminates channel estimation and alters the legitimate beamforming 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 joint 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. The proposed mitigation approach is illustrated via simulations.**

I. Introduction

Consider a three-node time-division duplex (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 channel state information (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 Eve attacks the channel training phase by transmitting the same pilot sequence during the training phase, 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 [1], [2].

This issue of pilot contamination attack was first noted in [1] where the focus is on enhancing eavesdropper's performance. Several approaches are discussed in [2]–[5] for detection of the attack assuming a TDMA uplink requiring separate time slots for each user Bob. In [6] an SDMA uplink was considered to allow for simultaneous transmission of training from Bobs. In [7], the approach of [4] was augmented with joint estimation of both legitimate receiver and eavesdropper channels, and secure beamforming, to mitigate the effects of pilot spoofing. In the set-up of [7], while the legitimate user superimposes a random sequence on its training sequence, the spoofer does not. In this paper we consider the scenario where the spoofer also could superimpose its own random sequence on the training sequence, before spoofing the legitimate user.

Approaches of [3], [5] require a separate secure channel from Alice-to-Bob (two-way training) to work. We only need one-way reverse training in this paper. Refs. [2], [4], [6] deal only with attack detection, not its mitigation. Unlike [7], we allow the spoofer to superimpose its own random sequence on the training sequence.

This work was supported by the National Science Foundation under Grant ECCS-1651133.

II. System Model and Background

We follow the system model of [2]–[5]. Let $s_t(n)$, $1 \le n \le T$, denote the training sequence of length T time samples. Consider a flat Rayleigh fading environment with Bob-to-Alice channel denoted as $\mathbf{h}_B = \sqrt{d_B}\,\tilde{\mathbf{h}}_B \in \mathbb{C}^{N_r \times 1}$ and Eve-to-Alice channel denoted as $\mathbf{h}_E = \sqrt{d_E}\,\tilde{\mathbf{h}}_E \in \mathbb{C}^{N_r \times 1}$, where real scalars d_B and d_E represent respective path loss attenuations 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 contamination 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).

[4] addresses the problem: how to detect Eve's attack based only on the knowledge of $s_t(n)$ and $\mathbf{y}(n)$. [4] allocates a fraction β of the training power P_B at Bob 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: BPSK or QPSK, e.g.) to be transmitted by Bob along with (superimposed on) $s_t(n)$. That is, instead of $\sqrt{P_B}s_t(n)$, Bob transmits ($0 \le \beta < 1,\ \ n = 1, 2, \cdots, T$)

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

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 [4], [7] one has the following two hypotheses $\mathcal{H}_0$ (no attack) and $\mathcal{H}_1$ (attack present) for the received signal at Alice:

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

In this paper, 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

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

That is, Eve spoofs Bob more closely than in [4], [7].

A. Approach of [4]

It is based on the model (4). Define the correlation matrix of measurements as $(i = 0, 1)$ $\mathbf{R}_{y,i} = T^{-1} \sum_{n=1}^{T} \mathbb{E}\left\{ \mathbf{y}(n)\mathbf{y}^H(n) \,\middle|\, \mathcal{H}_i \right\}$ and the correlation matrix of source signals as $(i = 0, 1)$ $\mathbf{R}_{s,i} = T^{-1} \sum_{n=1}^{T} \mathbb{E}\left\{ [\mathbf{y}(n) - \mathbf{v}(n)][\mathbf{y}(n) - \mathbf{v}(n)]^H \,\middle|\, \mathcal{H}_i \right\}$. Then we have $\mathbf{R}_{y,i} = \mathbf{R}_{s,i} + \sigma_v^2 \mathbf{I}_{N_r}$, $i = 0, 1$. It is shown in [4] 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 $\text{rank}(\mathbf{R}_{s,1}) = 1$. [4] exploits the MDL estimator of the signal subspace dimension ([8]) based on the eigenvalues of the estimated data correlation matrix to detect spoofing attack; it does not address attack mitigation. 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.

B. Approach of [7]

If the MDL method indicates presence of attack, Alice proceeds to jointly estimate the channels to Bob and Eve. Consider a periodic training sequence $s_t(n)$ with period P and $T = n_b P$ for some integer $n_b \geq 1$ (we can have $n_b = 1$). Stack P consecutive samples of ℓth component $y_\ell(n)$ of $\mathbf{y}(n)$ into a column:

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

Define $\mathbf{v}^\ell(m)$ from $v_\ell(n)$, the ℓth component $\mathbf{v}(n)$ in a similar fashion. Let $\check{\mathbf{s}}_t = [s_t(1) \ s_t(2) \ \cdots \ s_t(P)]^\top$ and $\check{\mathbf{s}}_B(m) = [s_B(1 + (m-1)P) \ \cdots \ s_B(P + (m-1)P)]^\top$. Then in the presence of self-contamination and eavesdropper, we have

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

where $h_{B,\ell}$ is the ℓth component of $\mathbf{h}_B$, and similarly for $h_{E,\ell}$. Let $\mathcal{P}_{\check{\mathbf{s}}_t}^{\perp}$ = projection orthogonal to the subspace spanned by $\check{\mathbf{s}}_t$. Then $\mathcal{P}_{\check{\mathbf{s}}_t}^{\perp} \mathbf{y}^\ell(m)$ has no contribution from training $s_t(n)$. "Reshape" $\mathcal{P}_{\check{\mathbf{s}}_t}^{\perp} \mathbf{y}^\ell(m)$ into a row vector along time and put all components ℓs together. Then the so "projected" $\mathbf{y}(n)$, denoted by $\tilde{\mathbf{y}}(n)$, 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 if model (4) holds true. Omitting the details, we have

$$\tilde{\mathbf{y}}(n) = \sqrt{P_B \beta}\, \mathbf{h}_B \tilde{s}_B(n) + \tilde{\mathbf{v}}(n) \tag{7}$$

where $n = 1, 2, \cdots, n_b(P-1)$, $\{\tilde{\mathbf{v}}(n)\}$ is i.i.d. zero-mean complex Gaussian with covariance $\sigma_v^2 \mathbf{I}_{P-1}$ and similarly $\tilde{s}_B(n)$ is uncorrelated zero-mean sequence with $\mathbb{E}\{|\tilde{s}_B(n)|^2\}$ not a function of n (however, $\tilde{s}_B(n)$ is not i.i.d.). In [7], (7) is used to estimate $\mathbf{h}_B$, which in turn, is used to estimate $\mathbf{h}_E$ after estimating the composite channel $\mathbf{h}_c := \sqrt{P_B(1-\beta)}\, \mathbf{h}_B + \sqrt{P_E}\, \mathbf{h}_E$ using the training sequence $s_t(n)$ and least-squares.

III. Proposed Approach based on Model (5)

If (5) is true, then (7) becomes

$$\tilde{\mathbf{y}}(n) = \sqrt{P_B \beta}\, \mathbf{h}_B \tilde{s}_B(n) + \sqrt{P_E \beta_2}\, \mathbf{h}_E \tilde{s}_E(n) + \tilde{\mathbf{v}}(n). \tag{8}$$

While the signal subspace of (7) is of rank 1, the signal subspace of (8) is of rank 2. The MDL criterion applied to $\{\tilde{\mathbf{y}}(n)\}$ will reveal its signal subspace rank. If this rank is 1, the approach of [7] suffices. If this rank is 2, then (5) is true, and the approach of [7] will fail. (8) represents a mixture of two non-Gaussian signals in white Gaussian noise, but the non-Gaussian signals are not i.i.d., hence, the standard approaches for unmixing using higher-order statistics (e.g., kurtosis) [12], [13] do not apply.

Under model (5), let

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

Let us estimate $\mathbf{h}_c$ using the training sequence $s_t(n)$ and least-squares, as $\hat{\mathbf{h}}_c = \frac{1}{T} \sum_{n=1}^{T} \mathbf{y}(n) s_t^*(n)$. Define

$$\begin{aligned} \check{\mathbf{y}}(n) = &\mathbf{y}(n) - \hat{\mathbf{h}}_c s_t(n) \\ \approx &\sqrt{P_B \beta}\, \mathbf{h}_B s_B(n) + \sqrt{P_E \beta_2}\, \mathbf{h}_E s_E(n) + \mathbf{v}(n). \end{aligned} \tag{9}$$

Now we can apply higher-order statistics-based approaches to unmix and estimate (scaled) $\mathbf{h}_B$ and $\mathbf{h}_E$. We use the RobustICA algorithm of [12] that uses kurtosis of "unmixed" measurements, after whitening $\check{\mathbf{y}}(n)$. Details follow.

Let $\mathbf{R}_{\check{y}} = \frac{1}{T} \sum_{n=1}^{T} \mathbb{E}\{\check{\mathbf{y}}(n)\check{\mathbf{y}}^H(n)\}$ and $\widehat{\mathbf{R}}_{\check{y}} = \frac{1}{T} \sum_{n=1}^{T} \check{\mathbf{y}}(n)\check{\mathbf{y}}^H(n)$, where $\widehat{\mathbf{R}}_{\check{y}}$ is a consistent estimator of the correlation matrix $\mathbf{R}_{\check{y}}$ of $\{\check{\mathbf{y}}(n)\}$. Consider EVD of $\widehat{\mathbf{R}}_{\check{y}}$ to obtain

$$\widehat{\mathbf{R}}_{\check{y}} = \hat{\mathbf{U}} \hat{\Sigma} \hat{\mathbf{U}}^H = [\hat{\mathbf{U}}_1 \ \hat{\mathbf{U}}_2] \begin{bmatrix} \hat{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \hat{\Sigma}_2 \end{bmatrix} [\hat{\mathbf{U}}_1 \ \hat{\mathbf{U}}_2]^H \tag{10}$$

where $\hat{\Sigma}$ is a $N_r \times N_r$ diagonal matrix with the eigenvalues of $\widehat{\mathbf{R}}_{\check{y}}$ arranged in decreasing order of magnitude, columns of $\hat{\mathbf{U}}$ are the corresponding eigenvectors, and $\hat{\mathbf{U}}_1$ is $N_r \times 2$. With reference to (9), define a channel matrix $\mathbf{H}_d \in \mathbb{C}^{N_r \times 2}$ as

$$\mathbf{H}_d = [\sqrt{P_B \beta}\, \mathbf{h}_B \ \sqrt{P_E \beta_2}\, \mathbf{h}_E]. \tag{11}$$

Then we can rewrite (9) as

$$\check{\mathbf{y}}(n) = \mathbf{H}_d \mathbf{s}(n) + \tilde{\mathbf{v}}(n), \quad \mathbf{s}(n) = [s_B(n) \ s_E(n)]^T. \tag{12}$$

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

$$\mathbf{R}_{\check{\mathbf{y}}} = \mathbf{U}\Sigma\mathbf{U}^H = [\mathbf{U}_1\ \mathbf{U}_2] \begin{bmatrix} \Sigma_1 & \mathbf{0} \\ \mathbf{0} & \Sigma_2 \end{bmatrix} [\mathbf{U}_1\ \mathbf{U}_2]^H \tag{13}$$

$$= \mathbf{H}_d\mathbf{H}_d^H + \sigma_v^2\mathbf{I}_{N_r} \tag{14}$$

where $\mathbf{U}$, Σ, etc. in (13) are the true counterparts of the estimated $\hat{\mathbf{U}}$, $\hat{\Sigma}$, etc. in (10).

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

$$\mathbf{x}(n) = \mathbf{U}_1^H\left(\mathbf{H}_d\mathbf{s}(n) + \tilde{\mathbf{v}}(n)\right) = \tilde{\mathbf{H}}_d\mathbf{s}(n) + \check{\mathbf{v}}(n) \tag{15}$$

where $\tilde{\mathbf{H}}_d \in \mathbb{C}^{2\times 2}$, $\tilde{\mathbf{H}}_d = \mathbf{U}_1^H\mathbf{H}_d$, and $\check{\mathbf{v}}(n) = \mathbf{U}_1^H\tilde{\mathbf{v}}(n) \in \mathbb{C}^2$. 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 [12] to yield an estimate $\hat{\tilde{\mathbf{H}}}_d$ of $\tilde{\mathbf{H}}_d$ using $\mathbf{x}(n)$. One obtains, for some $\theta_i s$,

$$\hat{\tilde{\mathbf{H}}}_d \approx \tilde{\mathbf{H}}_d\mathcal{P}\Gamma_\theta, \quad \Gamma_\theta = \mathrm{diag}\{e^{j\theta_i},\ i = 1, 2\} \tag{16}$$

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 [13]), 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

$$\hat{\mathbf{H}}_d = \hat{\mathbf{U}}_1\hat{\tilde{\mathbf{H}}}_d = [\sqrt{P_B\beta}\,\hat{\mathbf{h}}_B\ \ \sqrt{P_E\beta_2}\,\hat{\mathbf{h}}_E] \approx \mathbf{H}_d\mathcal{P}\Gamma_\theta\,. \tag{17}$$

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

$$\mathbf{z}(n) = \mathbf{A}\mathbf{s}(n) \tag{18}$$

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), and the objective is to recover $\mathbf{s}(n)$ and estimate $\mathbf{A}$. Such problems have been addressed in [12], [13], 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 (15), we see that (18) corresponds to (15) 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 [12], $\mathbf{w}$ is picked to maximize $|\gamma_4|$, where the kurtosis (normalized 4th cumulant) γ_4 of $e(n)$ is given by

$$\gamma_4 = \frac{\mathbb{E}\{|e(n)|^4\} - 2\mathbb{E}\{|e(n)|^2\} - \mathbb{E}\{e^2(n)\}}{\left(\mathbb{E}\{|e(n)|^2\}\right)^2}\,. \tag{19}$$

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 $s_m(n)$, (18) 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. □

Remark 2: Note that order of the extracted sources in an ICA problem is unknown [12], [13]. 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 (9), we do not know if an estimated channel resulting from the application of the RobustICA algorithm of [12] (or any other unmixing approach), corresponds to that of Bob or of Eve. We need some additional information to resolve this ambiguity. If old 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 the superimposed random sequence of Bob has some information embedded in it regarding user identification, and Alice can extract this from decoded data, decoded using, for instance, matched filter beamforming based on estimated channel (see Sec. IV). In the absence of such information, Alice quits if MDL-based subspace rank determined from $\{\tilde{\mathbf{y}}(n)\}$ of (7) exceeds one. □

IV. Secure Beamforming

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

$$y_B(n) = \sqrt{P_A}\mathbf{h}_B^\top\mathbf{w}\,s_A(n) + v_B(n) \tag{20}$$

$$y_{AE}(n) = \sqrt{P_A}\mathbf{h}_E^\top\mathbf{w}\,s_A(n) + v_E(n), \tag{21}$$

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. For MF reception at Bob, Alice should pick $\mathbf{w}$ as $\mathbf{h}_B^*/\|\mathbf{h}_B\|$ if $\mathbf{h}_B$ is known [9], [10], but instead uses the estimated channel to pick

$$\mathbf{w}_* = \widehat{\mathbf{h}}_B^*/\|\widehat{\mathbf{h}}_B\|. \tag{22}$$

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.

The SNRs at Bob and Eve, respectively, are $\mathrm{SNR}_B = P_A|\mathbf{h}_B^\top\mathbf{w}_*|^2/\sigma_B^2$, $\mathrm{SNR}_E = P_A|\mathbf{h}_E^\top\mathbf{w}_*|^2/\sigma_E^2$. 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

$$R_{B,sec} = \max(R_B - R_E, 0)\,. \tag{23}$$

In the presence of Eve with channel $\mathbf{h}_E$, the beamformer $\mathbf{w}$ may be picked to maximize $R_{B,sec}$. By [11, 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

$$\left(\mathbf{I}_{N_r} + \mathbf{h}_B^*\mathbf{h}_B^\top/\sigma_B^2,\ \mathbf{I}_{N_r} + \mathbf{h}_E^*\mathbf{h}_E^\top/\sigma_E^2\right). \tag{24}$$

We will use (24) with true channels replaced with their estimated counterparts.

V. Simulation Example

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 her random sequence $s_E(n)$. 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. Figs. 1-2 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. 1 applies to model (4), i.e., $\beta_2 = 0$ in (5), and Fig. 2 applies to model (5) with $\beta_2 = 0$.

The secrecy rate results of precoding (i.e., matched filter beamforming (22) or (24)) are shown in Figs. 3-4, with the corresponding channel estimation phase-insensitive MSE (mean-square error) shown in Figs. 5-6 and 7-8 for Eve's and Bob's channels, respectively, all averaged over 5000 runs. If $\widehat{\mathbf{h}}_B$ is an estimate of $\mathbf{h}_B$, both normalized to unit norm, phase-insensitive MSE is given by [14]

$$2 - 2|\mathbf{h}_B^H \widehat{\mathbf{h}}_B| = \min_{\theta\in[0,2\pi]} \|\mathbf{h}_B - e^{j\theta}\widehat{\mathbf{h}}_B\|^2.$$

The precoders/beamformers do not depend upon any phase rotation θ, hence, the use of phase-insensitive MSE to evaluate channel estimation errors in our context (as in [14] in a different problem).

It is seen from Fig. 3 that when $\beta_2 = 0$ (Eve sends only training), secure beamforming (precoding) yields a secrecy rate performance as a function of P_E that is almost invariant to the presence/absence of pilot spoofing attack. However, as seen in Fig. 4, when Eve also sends a random sequence, she reveals more of herself, resulting in better secrecy rate for Bob with increasing P_E. As seen in Fig. 6, $\beta_2 > 0$ yields better Eve's channel estimates with increasing P_E compared to the $\beta_2 = 0$ case (Fig. 5), hence, better nulling by the beamformer of Alice along Eve's direction. On the other hand, Bob's channel estimate accuracy is relatively unaffected in the two cases, as seen in Figs. 7-8.

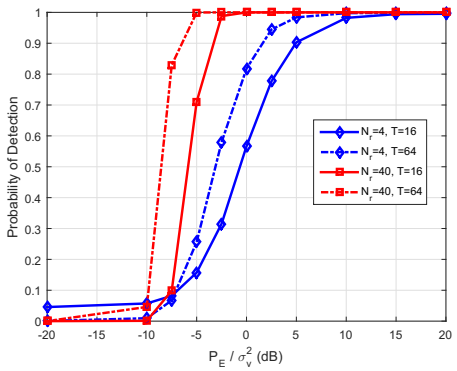

Fig. 1. *Probability of attack detection using MDL, as a function of Eve's power P_E relative to noise power σ_v^2 when Bob's power is fixed at $P_B/\sigma_v^2 = 10dB$. $\mathbf{y}(n)$ follows (4), β=0.4, $\beta_2 = 0$.*

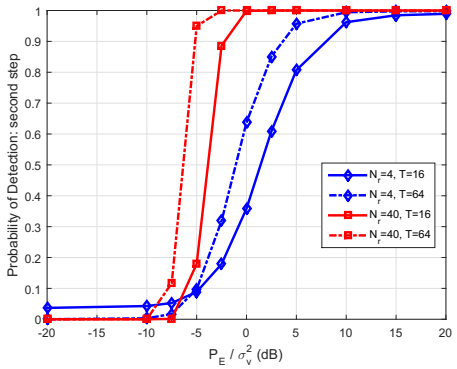

Fig. 2. *As in Fig. 1 except $\mathbf{y}(n)$ follows (5) with $\beta = \beta_2$=0.4 .*

References

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

[2] 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)*, pp. 13-18, London, UK, Sept. 8-11, 2013.

[3] 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. Information Forensics & Security*, vol. 10, pp. 932-940, May 2015.

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

[5] 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. Information Forensics & Security*, vol. 11, pp. 1017-1026, May 2016.

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

[7] J.K. Tugnait, "On mitigation of pilot spoofing attack," in *Proc. 2017 IEEE Intern. Conf. Acoustics, Speech & Signal Processing (ICASSP 2017)*, New Orleans, Louisiana, March 5-9, 2017.

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

[9] 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 Proc.*, vol. 8, no. 5, pp. 742-758, Oct. 2014.

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

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

[12] 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.

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

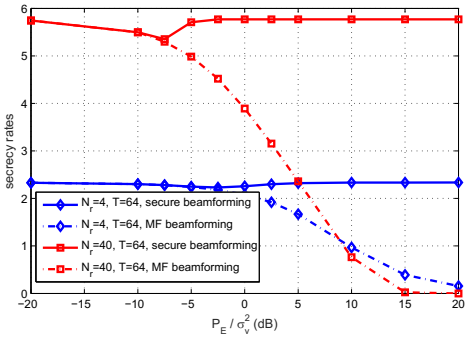

Fig. 3. *Secrecy rate (bps/Hz) at Bob using the beamformers based on (22) or (24), as a function of Eve's power P_E. All parameters as for Fig. 1. The label "precoding" refers to matched filter beamforming based on (24); "no precoding" means ones uses (22) with Eve ignored in channel estimation. $P_A = 1$, $\sigma_B^2 = \sigma_E^2 = 0.1$*

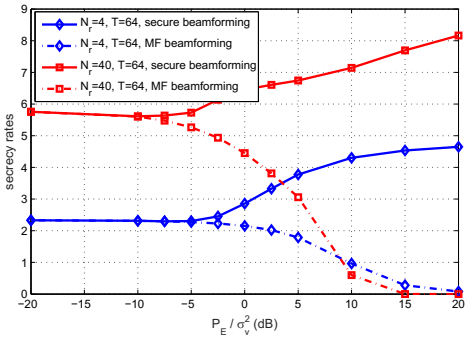

Fig. 4. *As in Fig. 3 except Eve also superimposes a random sequence, i.e., $\mathbf{y}(n)$ follows (5) with $\beta = \beta_2$=0.4 .*

[14] 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.

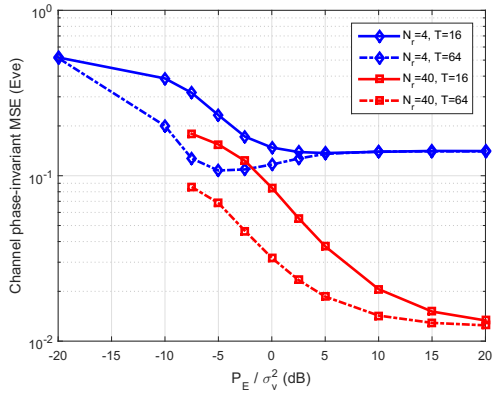

Fig. 5. *Channel phase-insensitive MSE for Eve's channel as a function of Eve's power P_E. All parameters as for Fig. 1.*

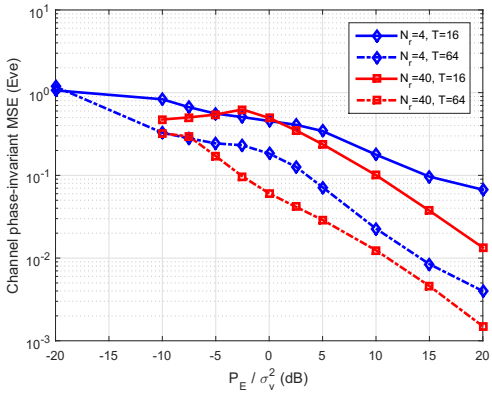

Fig. 6. *As in Fig. 5 except Eve also superimposes a random sequence.*

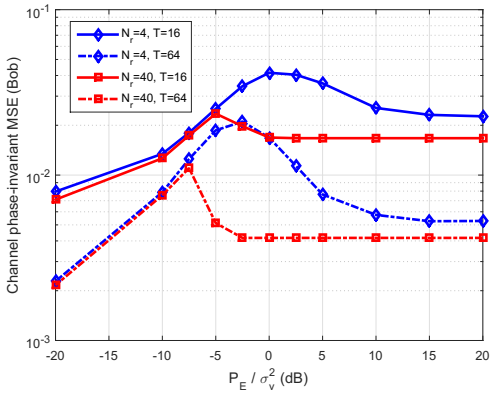

Fig. 7. *Channel phase-insensitive MSE for Bob's channel as a function of Eve's power P_E. All parameters as for Fig. 1.*

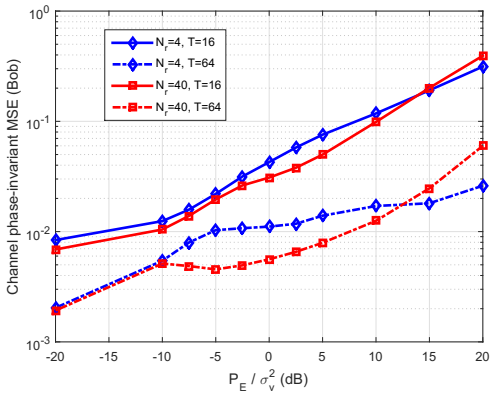

Fig. 8. *As in Fig. 7 except Eve also superimposes a random sequence.*