Mitigation of Pilot Spoofing Attack in Frequency Selective Channels

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 also sends identical training (pilot) signal as that of the legitimate receiver. This contaminates channel estimation and alters the legitimate beamformimg design, facilitating eavesdropping. Most of past approaches to pilot spoofing detection are limited to flat fading channels. A recent approach proposed superimposing a random sequence on the training sequence at the legitimate receiver for detection of pilot spoofing attack over frequency selective channels, with unknown channels and channel lengths, except that an upperbound on the number of channel taps is assumed to be known. In this paper we augment this approach with joint estimation of both legitimate receiver and eavesdropper channels, and secure time-reversal precoding, to mitigate the effects of pilot spoofing. 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 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, synchronized with Bob's training. Then the CSI estimated by Alice 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].

Several types of eavesdropping have been identified and analyzed in the literature [2]. 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 [3]. Such jamming attacks may occur during the training phase (pilot jamming) and/or in the data phase. 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 [1], [2], [4], 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. 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.

Almost all prior works on pilot spoofing detection [2], [4]–[9] deal with flat fading environments. The assumption of flat fading is fundamental to these cited papers, and their solutions will not work in frequency selective channels. In this paper we address frequency selective channels with unknown channels and channel lengths. In contrast, prior works such as [2], [4]–[9], assume that channels are 1-tap channels. Spoofing detection over frequency selective channels was recently addressed in [10]. In this paper we augment this approach with joint estimation of both legitimate receiver and eavesdropper channels, and secure time-reversal precoding, to mitigate the effects of pilot spoofing.

II. System Model

We consider an MISO (multiple-input single-output) system with a multi-antenna transmitter Alice equipped with N_r antennas, a single antenna legitimate user Bob, and an eavesdropper Eve. 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 been investigated in [4]–[6], except that instead of considering flat fading channels, we consider frequency selective channels.

Let $s_t(n)$, $1 \le n \le T$, denote the training sequence of length T time samples. Bob-to-Alice frequency selective channel impulse response is denoted as $\{\mathbf{h}_{B\ell}\}_{\ell=0}^{L_B-1}$ ($\mathbf{h}_{B\ell} \in \mathcal{C}^{N_r}$, L_B is the Bob's channel length (number of taps)), and Eve-to-Alice channel is denoted as $\{\mathbf{h}_{E\ell}\}_{\ell=0}^{L_E-1}$ ($\mathbf{h}_{E\ell} \in \mathcal{C}^{N_r}$, L_E is the Eve's channel length), where the impulse responses include both large-scale and small-scale fading effects. Let P_B and P_E

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

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

$$\mathbf{y}(n) = \sqrt{P_B} \sum_{\ell=0}^{L_B-1} \mathbf{h}_{B\ell} s_t(n-\ell) + \mathbf{v}(n) \in \mathcal{C}^{N_r} \quad (1)$$

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 pilot, the received signal at Alice during the training phase is ($\bar{L} = \max(L_B, L_E)$)

$$\mathbf{y}(n) = \sum_{\ell=0}^{\bar{L}-1} \left(\sqrt{P_B}\, \mathbf{h}_{B\ell} + \sqrt{P_E}\, \mathbf{h}_{E\ell} \right) s_t(n-\ell) + \mathbf{v}(n) \quad (2)$$

where $\mathbf{h}_{B\ell} = 0$ for $\ell \geq L_B$ and $\mathbf{h}_{E\ell} = 0$ for $\ell \geq L_E$. In case of Eve's attack, based on (2), Alice would estimate $\sqrt{P_B}\, \mathbf{h}_{B\ell} + \sqrt{P_E}\, \mathbf{h}_{E\ell}$, $\ell = 0, 1, \cdots$, as Bob-to-Alice channel, instead of $\sqrt{P_B}\, \mathbf{h}_{B\ell}$ based on (1).

A. Self-contamination at Bob

How to detect Eve's attack based only on the knowledge of $s_t(n)$ and $\mathbf{y}(n)$, is addressed in [5] for flat fading channels, 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: 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 ($0 \leq \beta < 1, \;\; n = 1, 2, \cdots, T$)

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

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 this case, in the absence of any transmission from Eve, the received signal at Alice during the training phase is given by

$$\mathbf{y}(n) = \mathbf{x}_0(n) + \mathbf{v}(n), \;\; \mathbf{x}_0(n) = \sum_{\ell=0}^{L_B-1} \mathbf{h}_{B\ell} \tilde{s}_B(n-\ell)\,. \quad (4)$$

When Eve also transmits, we have

$$\mathbf{y}(n) = \mathbf{x}_1(n) + \mathbf{v}(n) \quad (5)$$

where

$$\mathbf{x}_1(n) = \sum_{\ell=0}^{L_B-1} \mathbf{h}_{B\ell} \tilde{s}_B(n-\ell) + \sqrt{P_E} \sum_{\ell=0}^{L_E-1} \mathbf{h}_{E\ell} s_t(n-\ell). \quad (6)$$

In [10] we extended the self-contamination approach of [5] to apply to frequency selective channels. Let $L_m \geq \bar{L} = \max(L_B, L_E)$ and $T_m = T - L_m + 1$. We do not assume knowledge of L_B or L_E, but an upperbound L_m on them is assumed to be known to Alice. Define the $L_m \times T_m$ matrix

$$\mathbb{U} = \begin{bmatrix} s_t(L_m) & s_t(L_m+1) & \cdots & s_t(T) \\ s_t(L_m-1) & s_t(L_m) & \cdots & s_t(T-1) \\ \vdots & \vdots & \ddots & \vdots \\ s_t(1) & s_t(2) & \cdots & s_t(T-L_m+1) \end{bmatrix} \quad (7)$$

We assume that $\{s_t(n)\}$ is such that $\rho(\mathbb{U}) = L_m$. It then follows that $\rho(\mathbb{U}\mathbb{U}^H) = L_m$. This is the persistence of excitation condition of order L_m [11, Def. 10.1], which is necessary and sufficient for unique estimation of channel tap gains (for number of taps $\leq L_m$) using the method of least squares.

III. Attack Detection

Now we have the following two hypotheses $\mathcal{H}_0$ (no attack) and $\mathcal{H}_1$ (attack present) for the received signal at Alice:

$$\begin{array}{ll} \mathcal{H}_0: & \mathbf{y}(n) = \mathbf{x}_0(n) + \mathbf{v}(n) \\ \mathcal{H}_1: & \mathbf{y}(n) = \mathbf{x}_1(n) + \mathbf{v}(n) \end{array}, n = 1, 2, \cdots, T. \quad (8)$$

A. Signal Subspace Dimension

Define the correlation matrices $\mathbf{R}_{y,i}$ and $\mathbf{R}_{x,i}$ of measurements and signals, respectively, as ($i = 0, 1$)

$$\mathbf{R}_{y,i} = T_m^{-1} \sum_{n=L_m}^{T} \mathbb{E}\left\{ \mathbf{y}(n)\mathbf{y}^H(n) \,\middle|\, \mathcal{H}_i \right\}, \quad (9)$$

$$\mathbf{R}_{x,i} = T_m^{-1} \sum_{n=L_m}^{T} \mathbb{E}\left\{ \mathbf{x}_i(n)\mathbf{x}_i^H(n) \,\middle|\, \mathcal{H}_i \right\}. \quad (10)$$

Then we have $\mathbf{R}_{y,i} = \mathbf{R}_{x,i} + \sigma_v^2 \mathbf{I}_{N_r}, \;\; i = 0, 1$. It is shown in [10] that $\mathrm{rank}(\mathbf{R}_{x,0}) = L_B$ w.p.1 if $N_r \geq L_B$, and $\mathrm{rank}(\mathbf{R}_{x,1}) = L_B + L_E$ w.p.1 if $N_r \geq L_B + L_E$.

Thus, the ranks of the signal correlation matrix under the two hypotheses are different. Alice does not know the true values of L_B and L_E, only an upperbound L_m on them. Lack of knowledge of L_B and L_E precludes use of the approach of [5] (also used in [8], [9]), which relies on the knowledge that $L_B = L_E = 1$, i.e., the channels are flat-fading (1-tap). [10] proposed an alternative approach to attack detection. Here we follow a similar approach, discussed next, which differs in details.

B. Attack Detection Approach

Regardless of the absence/presence of spoofer, we first estimate the channel $\mathbf{h}_{C\ell} = \sqrt{(1-\beta)P_B}\mathbf{h}_{B\ell} + \sqrt{P_E}\mathbf{h}_{E\ell}$ with known input $s_t(n)$ and noisy output $\mathbf{y}(n)$ using the method of least-squares. The solution $\hat{\mathbf{h}}_{C\ell}$ satisfies ($k = 0, 1, \cdots, L_m - 1$)

$$\sum_{\ell=0}^{L_m-1} r_s(\ell, k) \hat{\mathbf{h}}_{C\ell} = \frac{1}{T_m} \sum_{n=L_m}^{T} \mathbf{y}(n) s_t^*(n-k),$$

where $r_s(\ell,k)=\frac{1}{T_m}\sum_{n=L_m}^{T} s_t(n-\ell)s_t^*(n-k)$. Remove the training contribution from the received signal to define

$$\begin{aligned}\tilde{\mathbf{y}}(n) =&\mathbf{y}(n)-\sum_{\ell=0}^{L_m-1}\hat{\mathbf{h}}_{C\ell}s_t(n-\ell)\\ \approx&\sqrt{\beta P_B}\sum_{\ell=0}^{L_B-1}\mathbf{h}_{B\ell}s_B(n-\ell)+\mathbf{v}(n)=\tilde{\mathbf{x}}(n)+\mathbf{v}(n).\end{aligned} \tag{11}$$

In addition to (8), consider the nature of projected $\{\tilde{\mathbf{y}}(n)\}$ under the two hypotheses:

$$\begin{array}{ll}\mathcal{H}_0: & \tilde{\mathbf{y}}(n)=\tilde{\mathbf{x}}(n)+\mathbf{v}(n)\\ \mathcal{H}_1: & \tilde{\mathbf{y}}(n)=\tilde{\mathbf{x}}(n)+\mathbf{v}(n)\end{array},n=L_m,L_m+1,\cdots,T_m\,. \tag{12}$$

We see that under $\mathcal{H}_0$, the signal subspace rank of both $\{\mathbf{y}(n)\}$ and $\{\tilde{\mathbf{y}}(n)\}$ is L_B, whereas under $\mathcal{H}_1$, the signal subspace rank of $\{\mathbf{y}(n)\}$ is L_B+L_E while that of $\{\tilde{\mathbf{y}}(n)\}$ is L_B. Since the channel lengths L_B and L_E are not known, our proposed relies on estimating the signal subspace ranks of $\{\mathbf{y}(n)\}$ and $\{\tilde{\mathbf{y}}(n)\}$: if the two ranks are the same, there is no pilot spoofing, and if the two ranks are different, one declares presence of a pilot spoofing attack. In contrast, in the approach of [5] (also used in [8], [9]) applicable to flat fading channels, it is enough to check the signal subspace rank of $\{\mathbf{y}(n)\}$, which is 1 if there is no pilot spoofing, and is 2 in the presence of pilot spoofing.

Two different approaches for estimation of signal subspace rank given observations of signals in white Gaussian noise, were used in [10]: the minimum description length (MDL) source enumeration method ([12]–[14]), and the random matrix theory (RMT) based source enumeration approach of [15], [16]. Note also that model (12) used here is different from that in [10].

IV. Channel Estimation

A. Estimation of Bob's Channel

Now using (11), we apply the blind approach of [17] (the SIMO case, equalizer length of 5 taps, delay of 2) to estimate $\mathbf{h}_{B\ell}$ as $\hat{\mathbf{h}}_{B\ell}=c\mathbf{h}_{B\ell}$, $\forall\ell$, up to a complex constant c. (Note that step 2 of Algorithm 1 of [17] was modified to extract "significant" principal eigenvectors of the data correlation matrix, instead of the number of principal eigenvectors stated in [17, Step 2, Alg. 1]. All eigenvalues smaller than $0.1\times$ the largest eigenvalue of the data correlation matrix in step 2 of Algorithm 1 of [17] were deemed to be insignificant, hence the corresponding eigenvectors were insignificant. The reason for this modification is the lack of knowledge of L_B in (11).) Also, [17] involves equalization and quantization of $s_B(n)$.

We will use a phase-insensitive mean-square error (MSE) measure to evaluate channel estimation errors; this has been used in [18] in a different context. If $\hat{\mathbf{h}}^{(B)}$ is an estimate of $\mathbf{h}^{(B)}=[\mathbf{h}_{B0}^\top\ \cdots\ \mathbf{h}_{B(L_m-1)}^\top]^\top$, both normalized to unit norm, phase-insensitive MSE in estimation of $\mathbf{h}^{(B)}$ is given by [18]

$$\min_{\theta\in[0,2\pi]}\|\mathbf{h}^{(B)}-e^{j\theta}\hat{\mathbf{h}}^{(B)}\|^2=2-2|\mathbf{h}^{(B)\,H}\hat{\mathbf{h}}^{(B)}|. \tag{13}$$

Correct scaling of $\hat{\mathbf{h}}^{(B)}$ is possible along the lines of [9] but as applied to frequency-selective channels, by using equalized/quantized $s_B(n)$.

B. Estimation of Eve's Channel

If the detector indicates the presence of Eve, we also estimate Eve's channel. Here we need $\hat{\mathbf{h}}^{(B)}$ with proper scale. Then $\hat{\mathbf{h}}_{E\ell}=\hat{\mathbf{h}}_{C\ell}-\hat{\mathbf{h}}_{B\ell}$.

V. Time-Reversal Precoding at Alice

We use time-reversal beamforming [19] at Alice from transmission to Bob. In the absence of spoofing, Alice designs the precoder based on estimated Bob's channel. If spoofing is present, Alice designs a constrained time-reversal precoder to maximize SNR at Bob while placing a null toward Eve (using Eve's estimated channel) at several "time lags." Let $\{s_A(n)\}$, $\mathbb{E}\{|s_A(n)|^2\}=1$, denote the scalar information sequence of Alice intended for Bob. Alice designs a time-revsersal precoder with impulse response $\mathbf{w}_\ell\in\mathbb{C}^{N_r}$, $0\le\ell\le L_m-1$, and transmits $\sqrt{P_A}\sum_{\ell=0}^{L_m-1}\mathbf{w}_\ell\,s_A(n-\ell)=\sqrt{P_A}\mathbf{w}_n\otimes s_A(n)$ where P_A is the transmit power, $\otimes$ denotes convolution and $\{\mathbf{w}_\ell\}$ is normalized to unit norm. The received signals at Bob and Eve are given, respectively, by

$$y_B(n)=\sqrt{P_A}\mathbf{h}_{Bn}^\top\otimes\mathbf{w}_n\otimes s_A(n)+v_B(n) \tag{14}$$

$$y_{AE}(n)=\sqrt{P_A}\mathbf{h}_{En}^\top\otimes\mathbf{w}_n\otimes s_A(n)+v_E(n), \tag{15}$$

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.

In the absence of Eve (or, Eve is not detected), to maximize $|g_B(\ell)|$ at $\ell=L_m-1$, ($g_B(\ell)=\mathbf{h}_{B\ell}^\top\otimes\mathbf{w}_\ell$), matched filter reception at Bob yields $\mathbf{w}_\ell=\mathbf{h}_{B(L_m-1-\ell)}^*$, $0\le\ell\le L_m-1$. When Eve is present, the precoder is designed to maximize $|g_B(\ell)|$ at $\ell=L_m-1$ subject to $g_E(\ell)=0\ \forall\ell$ where $g_E(\ell)=\mathbf{h}_{E\ell}^\top\otimes\mathbf{w}_\ell$. For channel and precoder lengths not exceeding L_m, we need to consider $g_E(\ell)$ for $0\le\ell\le 2L_m-2$. Define

$$\bar{\mathbf{w}}=[\mathbf{w}_0^\top\ \mathbf{w}_1^\top\ \cdots\ \mathbf{w}_{L_m-1}^\top]^\top, \tag{16}$$

$$\mathcal{H}_E=\begin{bmatrix}\mathbf{h}_{E0}^\top & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{h}_{E1}^\top & \mathbf{h}_{E0}^\top & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{h}_{E(L_m-1)}^\top & \mathbf{h}_{E(L_m-2)}^\top & \cdots & \mathbf{h}_{E0}^\top\\ \mathbf{0} & \mathbf{h}_{E(L_m-1)}^\top & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \cdots & \cdots & \mathbf{h}_{E(L_m-1)}^\top\end{bmatrix}, \tag{17}$$

$$\bar{\mathbf{h}}_B=[\mathbf{h}_{B(L_m-1)}^\top\ \mathbf{h}_{B(L_m-2)}^\top\ \cdots\ \mathbf{h}_{B0}^\top]^\top. \tag{18}$$

Then $g_E(\ell)=0$ for $0\le\ell\le 2L_m-2$ is equivalent to $\mathcal{H}_E\bar{\mathbf{w}}=\mathbf{0}$.

This leads to the optimization problem

$$\max_{\bar{\mathbf{w}}}\ |\bar{\mathbf{h}}_B^\top\bar{\mathbf{w}}|\quad\text{subject to}\ \ \mathcal{H}_E\bar{\mathbf{w}}=\mathbf{0},\ \|\bar{\mathbf{w}}\|=1. \tag{19}$$

The constraint $\mathcal{H}_E\bar{\mathbf{w}} = \mathbf{0}$ implies that $\bar{\mathbf{w}}$ lies in a subspace orthogonal to that spanned by $\mathcal{H}_E^H$, i.e., for some $\bar{\mathbf{w}}_0$, with $\mathcal{P}^{\perp}_{\mathcal{H}_E^H}$ denoting projection orthogonal to $\mathcal{H}_E^H$,

$$\bar{\mathbf{w}} = \mathcal{P}^{\perp}_{\mathcal{H}_E^H}\bar{\mathbf{w}}_0 = \left(\mathbf{I}_{N_r L_m} - \mathcal{H}_E^H(\mathcal{H}_E\mathcal{H}_E^H)^{-1}\mathcal{H}_E\right)\bar{\mathbf{w}}_0. \quad (20)$$

With $\tilde{\bar{\mathbf{h}}}_B := (\mathcal{P}^{\perp}_{\mathcal{H}_E^H})^{\top}\bar{\mathbf{h}}_B$, $|\bar{\mathbf{h}}_B^{\top}\bar{\mathbf{w}}| = |\tilde{\bar{\mathbf{h}}}_B^{\top}\bar{\mathbf{w}}_0|$ is maximized w.r.t. $\bar{\mathbf{w}}_0$ by an MF solution $\bar{\mathbf{w}}_{0*} = c\tilde{\bar{\mathbf{h}}}_B^*$ for some nonzero constant c. Since $\mathcal{P}^{\perp}_{\mathcal{H}_E^H}$ is a projection operator satisfying $\mathcal{P}^{\perp}_{\mathcal{H}_E^H}(\mathcal{P}^{\perp}_{\mathcal{H}_E^H})^H = \mathcal{P}^{\perp}_{\mathcal{H}_E^H}$, in terms of $\bar{\mathbf{w}}$, we have $\bar{\mathbf{w}} = \mathcal{P}^{\perp}_{\mathcal{H}_E^H}\bar{\mathbf{w}}_{0*} = c\mathcal{P}^{\perp}_{\mathcal{H}_E^H}\bar{\mathbf{h}}_B^*$, where c is picked to set $\|\bar{\mathbf{w}}\| = 1$. We note again that if Eve is not detected, we pick $\bar{\mathbf{w}} = \bar{\mathbf{h}}_B^*$.

In practice, we replace $\mathbf{h}_{B\ell}$ and $\mathbf{h}_{E\ell}$ with their estimates. Also, since L_m typically overestimates the true lengths L_B and L_E, we replace $(\mathcal{H}_E\mathcal{H}_E^H)^{-1}$ in (20) with its pseudo-inverse via SVD.

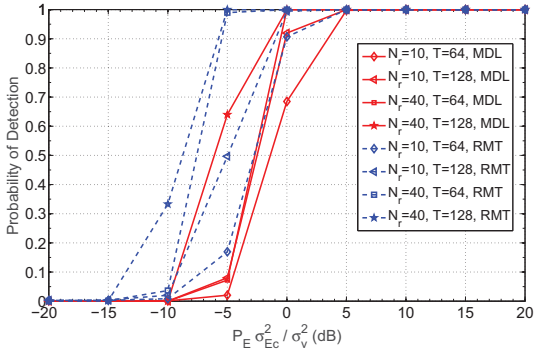

Fig. 1. Probability of attack detection as a function of Eve's power P_E relative to noise power σ_v^2 when Bob's power is fixed at $P_B\sigma_{Bc}^2/\sigma_v^2 = 10$dB, β=0.4 .

VI. Simulation Example

We consider frequency selective channels with $L_B = 3$, $L_E = 2$, both values unknown to Alice who uses the upper-bound $L_m = 4$, $\mathbf{h}_{B\ell} \sim \mathcal{N}_c(0, \sigma_{Bc}^2\mathbf{I}_{N_r})$, $\mathbf{h}_{E\ell} \sim \mathcal{N}_c(0, \sigma_{Ec}^2\mathbf{I}_{N_r})$, both channels have independent tap gains, and noise power σ_v^2, training power budget P_B at Bob is such that $P_B\sigma_{Bc}^2/\sigma_v^2 =$ 10dB, training power budget P_E at Eve is such that $P_E\sigma_{Ec}^2/\sigma_v^2$ varies from -20dB through 20dB, and fractional allocation β of training power at Bob to random sequence $s_B(n)$ is 0.4 . Bob and Eve have single antennas while Alice has $N_r = 10$ or 40 antennas ($\geq 2L_m$). The training sequence is a random binary sequence with T =64 or 128, and the random sequence $\{s_B(n)\}$ is i.i.d. QPSK. Fig. 1 shows our detection probability P_d results averaged over 5000 runs for both MDL and RMT (designed for false-alarm rate of 0.001) approaches. The performance improves with increasing T, N_r and Eve's power P_E, and RMT outperforms MDL. Fig. 2 shows phase-insensitive MSE in Bob's channel estimation. The curves labeled "blind" are based on the approach of [17], and the curves labeled "naive" ignore Eve's presence and use an iterative method for channel estimation (estimate channel using only training, equalize and quantize self-contamination, and then redo with training-plus-estimated $s_B(n)$ as pseudo-training). The blind result is invariant to Eve's power, since it is applied after canceling training contribution, hence Eve's contribution. The naive results work well for low P_E (as exptected), but rapidly deteriorate with increasing P_E.

The estimated Bob's channel can be used by Alice to implement a time-reversal matched-filter precoder (beamformer) [19] at Alice for transmission to Bob, as discussed in Sec. V. At Bob and Eve, respectively, we design linear MMSE equalizers with full knowledge of their respective channels and the beamformer at Alice, to evaluate possible performance limits. It is seen from Figs. 3-6 that when spoofing-aware channel estimates are used, Alice can frustrate Eve's eavesdropping with only a "small" (if any) deterioration in Bob's performance.

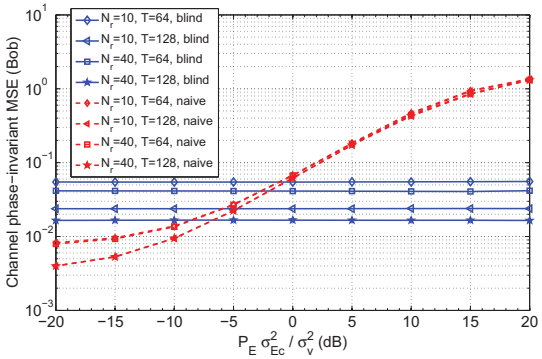

Fig. 2. Channel normalized MSE (13) for Bob's channel as a function of Eve's power P_E. All parameters as for Fig. 1.

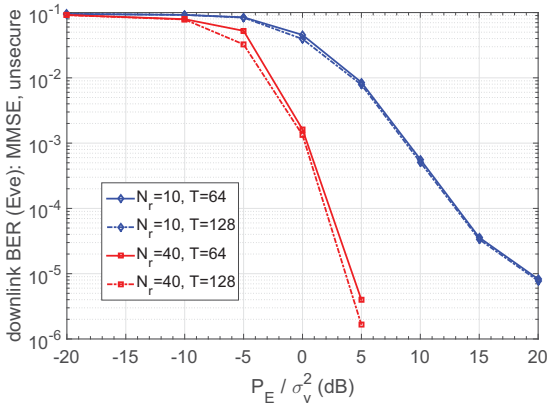

Fig. 3. BER at Eve when Alice uses naive estimate of Bob's channel for time-reversal beamforming design. Eve uses a linear MMSE equalizer with full knowledge of Alice-to-Eve channel and the beamformer at Alice.

VII. Conclusions

A novel approach to detection of pilot spoofing/contamination attack in a 3-node TDD system (legitimate source-destination pair Alice and Bob, and spoofer Eve)

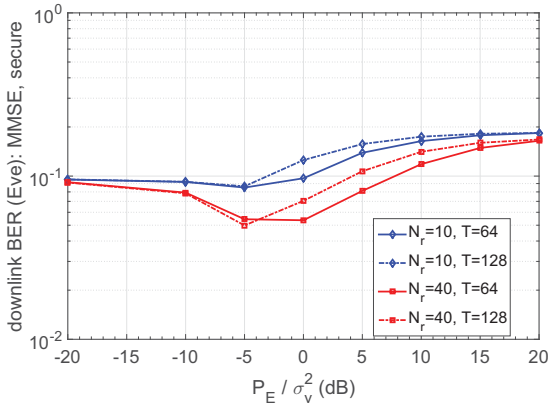

Fig. 4. BER at Eve when Alice uses spoofing-aware ("secure") estimate of Bob's channel, and Eve's estimated channel, for time-reversal beamforming design. Eve uses a linear MMSE equalizer with full knowledge of Alice-to-Eve channel and the beamformer at Alice.

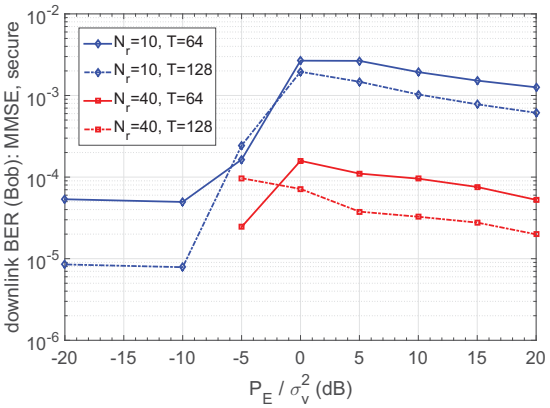

Fig. 6. BER at Bob when Alice uses spoofing-aware ("secure") estimate of Bob's channel, and Eve's estimated channel, for time-reversal beamforming design. Bob uses a linear MMSE equalizer with full knowledge of Alice-to-Bob channel and the beamformer at Alice.

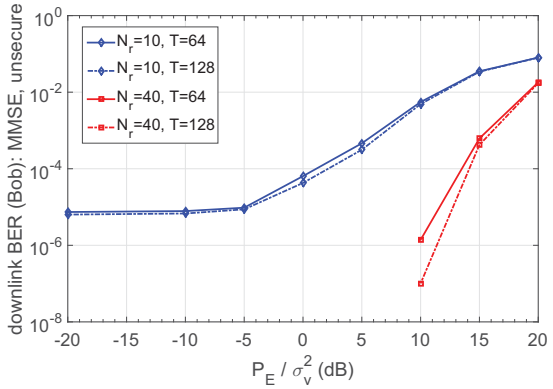

Fig. 5. BER at Bob when Alice uses naive estimate of Bob's channel for time-reversal beamforming design. Bob uses a linear MMSE equalizer with full knowledge of Alice-to-Bob channel and the beamformer at Alice.

was presented in [9] for frequency-selective channels, with unknown channels and channel lengths. In this paper we augmented this approach with joint estimation of both legitimate receiver and eavesdropper channels, and secure time-reversal precoding, to mitigate the effects of pilot spoofing. The proposed approach was illustrated by numerical examples and they show the efficacy of the proposed approach.

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 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

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

[9] J.K. Tugnait, "Pilot spoofing attack detection and countermeasure," *IEEE Trans. Commun.*, vol. 66, no. 5, pp. 2093-2106, May 2018.

[10] J.K. Tugnait, "Detection of pilot spoofing attack over frequency selective channels," in *Proc. 2018 IEEE Statistical Signal Processing Workshop (SSP)*, pp. 737-741, Freiburg, Germany, June 10-13, 2018.

[11] M. Verhaegen and V. Verdult, *Filtering and System Identification*. Cambridge, UK: Cambridge U. Press, 2007.

[12] 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.

[13] 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 Processing*, vol. 58, no. 1, pp. 452-457, Jan. 2010.

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

[15] 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.

[16] 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.

[17] I. Kacha, K. Abed-Meraim and A. Belouchrani, "Fast adaptive blind MMSE equalizer for multichannel FIR systems," *EURASIP J. Applied Signal Process.*, vol. 2006, Article ID 14827, pages 1-17, 2006.

[18] 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.

[19] T. Strohmer, M. Emami, J. Hansen, G. Papanicolaou and A.J. Paulraj, "Application of time-reversal with MMSE equalizer to UWB communications," in *Proc. IEEE Globecom 2004*, vol. 5, pp. 3123-3127, Nov. 2004.