Blind Channel Direction Separation Against Pilot Spoofing Attack in Massive MIMO System

Ruohan Cao, Tan F. Wong, Hui Gao, Dongqing Wang, and Yueming Lu

***Abstract*—This paper considers a pilot spoofing attack scenario in a massive MIMO system. A malicious user tries to disturb the channel estimation process by sending interference symbols to the basestation (BS) via the uplink. Another legitimate user counters by sending random symbols. The BS does not possess any partial channel state information (CSI) and distribution of symbols sent by malicious user *a priori*. For such scenario, this paper aims to separate the channel directions from the legitimate and malicious users to the BS, respectively. A blind channel separation algorithm based on estimating the characteristic function of the distribution of the signal space vector is proposed. Simulation results show that the proposed algorithm provides good channel separation performance in a typical massive MIMO system.**

***Index Terms*—Massive MIMO, pilot spoofing attack, blind channel separation.**

I. Introduction

Massive multiple-input multiple-output (MMIMO) systems [1], [2] exhibit excellent potentials for opposing passive eavesdropping attacks by using physical-layer security techniques [3]. Most of these physical-layer security techniques however rely on knowledge of channel state information (CSI), which is commonly estimated in the training phase of a MMIMO system. It is well known that MMIMO is vulnerable to active pilot spoofing attacks that aim to disturb the CSI estimation process [4]. This spoofing-attack vulnerability presents a weak spot for implementing physical-layer security techniques in MMIMO.

Signal processing methods have been recently proposed to counter pilot spoofing attacks in MMIMO systems. In particular, Refs. [4]–[13] propose attack detection methods that determine whether a pilot spoofing attack is conducted or not. Among these works, in [5] and [6], the BS performs attack detection by comparing the statistical properties of its observations with partial CSI known *a priori*. The proposed methods of [7]–[10] make the legitimate user send random pilot symbols to the BS. The randomness of the pilot symbols reduces the effect of pilot contamination, and allows the BS to detect the attack by determining the number of sources from which its observations come. With sightly difference, refs. [11] and [12] utilizes the randomness of data. Refs. [13] and [4] give detection methods performed by the legitimate user.

On the other hand, there are also efforts focusing on estimating the channels of the legitimate and malicious users [14]–[18]. The proposed methods all utilize different forms of *asymmetry* between the legitimate and malicious users. In [14] and [15], the BS could communicate with the legitimate user via a aided channel that the malicious user cannot access. In [16]–[18], the BS firstly uses independent component analysis (ICA) for separating channels, and then employ *asymmetry* to match these separated channels with users. To be more specific, in [16], it is assumed that some partial CSI (e.g., the path loss values) is known to the BS *a priori*, and that the *a priori* partial CSI of the legitimate user is different from that of the malicious user. The BS identifies channels by using the CSI difference. In [17] and [18], the *asymmetry* is based on the restriction that the legitimate user can send encrypted information to the BS while the malicious user cannot do that. We notice that for enabling the ICA separation method, the malicious user is assumed to send information following certain statistical distribution that is known to the BS and independent with symbols of legitimate user [16]–[18]. This assumption may not always true in practical scenarios.

Against this background, we focus on separating the channel directions from the BS to the legitimate user and the malicious user upon positive detection of such attacks has been made. In specific, we consider attack scenario in which:

1) No partial CSI is available to the BS *a priori*. Both the legitimate and malicious user channels are totally unknown to the BS.
2) The malicious user is free to impersonate the legitimate user. For instance, the malicious user may overhear the symbols sent by the legitimate user, and sends symbols according to its overheard signals. The employed distribution is unknown to the BS.

With channel direction separation, the BS can separately beamform to the legitimate user and the malicious user, and further employ some *asymmetry* configurations (e.g., higher layer authentication protocols) to distinguish between them.

We propose a blind channel separation algorithm in which the BS quantizes its observations, and obtains the empirical distribution of quantized observations for channel separation. We observe that this empirical distribution is determined by the channel directions and the distribution of the data symbols. As such, the channel directions can be extracted from the empirical distribution observed by the BS, and hence achieving

The financial support of the National Key Research and Development Program (Grant No. 2016YFB0800302), of the National Natural Science Foundation of China (NSFC) (Grant No. 61501046), and of the Fundamental Research Funds for the Central Universities of China (Grant No. 2482017RC01) is gratefully acknowledged.

R. Cao, H. Gao, D. Wang and Y. Lu are with the Key Laboratory of Trustworthy Distributed Computing and Service, Ministry of Education, and also the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications (BUPT), Beijing 100876, China (e-mail: {caoruohan, huigao, wangdongqing, ymlu}@bupt.edu.cn).

T. F. Wong is with the Department of Electrical and Computer Engineering, University of Florida, FL 32611, U.S.A. (e-mail: twong@ufl.edu).

ISBN 978-90-827970-1-5 © EURASIP 2018

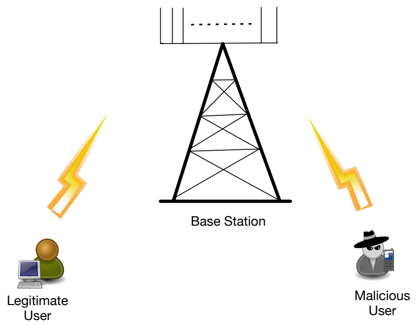

Fig. 1: System model: The BS is equipped with M antennas. The legitimate and malicious users have a single antenna each.

channel separation. In addition, the proposed algorithm is based on FFT, making it easy to be implemented in practice.

II. System Model

Assuming that pilot spoofing attack detection has been performed to validate the existence of a malicious user, we consider the system model depicted in Fig. 1. The BS is equipped with M antennas, while the legitimate and malicious users have a single antenna each. The uplink and the corresponding downlink channels satisfy reciprocity. We assume that the channel separation process is performed within the coherent time of the channels, and use the $M \times 1$ vectors $\boldsymbol{h}$ and $\boldsymbol{g}$ to specify the channels from the legitimate user and malicious user to the BS, respectively. In the uplink, the legitimate and malicious users send random symbols A and B at power P_1 and P_2 to the BS, respectively. We use P_A and P_B for denoting the stochastic distributions of random symbols A and B over finite alphabets $\mathcal{A}$ and $\mathcal{B}$, respectively. For arbitrary $\mathsf{a} \in \mathcal{A}$ and $\mathsf{b} \in \mathcal{B}$, we assume the joint distribution of A and B satisfies $P_{A,B}(\mathsf{a}, \mathsf{b}) > 0$ whenever $P_A(\mathsf{a}) > 0$ and $P_B(\mathsf{b}) > 0$. This assumption indicates A and B maybe dependent with each other, but B cannot be definitely determined by A. In practical, it corresponds to a fact that the malicious user sends B according to its overheard version of A. Nevertheless, the malicious user cannot get A exactly due to the channel fading and noise.

The received symbol of the BS is specified by

$$\boldsymbol{y} = \sqrt{P_1}\boldsymbol{h}A + \sqrt{P_2}\boldsymbol{g}B + \boldsymbol{w}, \tag{1}$$

where $\boldsymbol{w}$ is the noise vector, whose elements are i.i.d. circular-symmetric complex Gaussian (CSCG) random variables with zero mean and variance σ^2. Whenever needed, we use β_1 and β_2 for denoting the path losses of channels from the legitimate and malicious users to the BS, respectively. The BS does not know β_1, β_2, P_1, P_2, $\boldsymbol{h}$, $\boldsymbol{g}$ and P_B *a priori*.

Assuming the uplink channel described by (1) is used n times, within the coherent time of the channel, for the ith instant of use, (1) gives rise to

$$\boldsymbol{y}_i = \sqrt{P_1}\boldsymbol{h}A_i + \sqrt{P_2}\boldsymbol{g}B_i + \boldsymbol{w}_i, \tag{2}$$

for $i = 1, 2, \ldots, n$. Stacking the n equations in (2) into a matrix form, we obtain

$$\mathbf{Y} = [\boldsymbol{h}, \boldsymbol{g}] \begin{bmatrix} \mathbf{A}\sqrt{P_1} \\ \mathbf{B}\sqrt{P_2} \end{bmatrix} + \mathbf{W} \tag{3}$$

where $\mathbf{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n]$, $\mathbf{W} = [\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n]$, $\mathbf{A} = [A_1, A_2, \ldots, A_n]$, and $\mathbf{B} = [B_1, B_2, \ldots, B_n]$. Let $\boldsymbol{S}$ be the $M \times 2$ signal subspace matrix whose columns form an orthonormal basis that spans the column space of $[\boldsymbol{h}, \boldsymbol{g}]$. Then, we project $\mathbf{Y}$ onto the signal subspace and get from (3)

$$\mathbf{Z} = \frac{1}{\sqrt{M}}\boldsymbol{S}^T\mathbf{Y} = [\boldsymbol{z}_1', \boldsymbol{z}_2'] \begin{bmatrix} \mathbf{A}\sqrt{P_1} \\ \mathbf{B}\sqrt{P_2} \end{bmatrix} + \mathbf{N} \tag{4}$$

where $[\boldsymbol{z}_1', \boldsymbol{z}_2'] = \frac{1}{\sqrt{M}}\boldsymbol{S}^T[\boldsymbol{h}, \boldsymbol{g}]$, $\mathbf{N} = \frac{1}{\sqrt{M}}\boldsymbol{S}^T\mathbf{W}$, and the elements of $\mathbf{N}$ are i.i.d. Gaussian random variables with zero mean and variance $\frac{\sigma^2}{M}$. Clearly, $\mathbf{N}$ is independent of $[\boldsymbol{z}_1', \boldsymbol{z}_2'] \begin{bmatrix} \mathbf{A}\sqrt{P_1} \\ \mathbf{B}\sqrt{P_2} \end{bmatrix}$. It is argued in [19] that this independence and the Gaussianity of $\mathbf{N}$ imply that

$$\left[\hat{\boldsymbol{h}}, \hat{\boldsymbol{g}}\right] = \sqrt{M}\boldsymbol{S}[\boldsymbol{z}_1', \boldsymbol{z}_2'] \tag{5}$$

would be a reasonable estimator for the channel pair $[\boldsymbol{h}, \boldsymbol{g}]$ if $[\boldsymbol{z}_1', \boldsymbol{z}_2']$ could be found. In practice $\boldsymbol{S}$ is not known *a priori*, but can be estimated from the singular value decomposition $\mathbf{Y} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$ with orthogonal matrices $\boldsymbol{U} \in \mathbb{R}^{M \times M}$, $\boldsymbol{V} \in \mathbb{R}^{n \times n}$ and $M \times n$ diagonal singular value matrix $\boldsymbol{\Sigma}$. Then $\boldsymbol{S}$ can well approximated by the first two left singular vectors in $\boldsymbol{U}$ when M is large in the MMIMO system [19].

Since $\mathbf{A}$ and $\mathbf{B}$ may have the same distribution, $[\boldsymbol{z}_1', \boldsymbol{z}_2']$ cannot be uniquely determined from $\mathbf{Y}$ without knowledge of β_1, β_2, P_1, P_2, $\mathbf{A}$, and $\mathbf{B}$. As will be argued in the next section, it is however possible to estimate $\left[\frac{\boldsymbol{h}}{|\boldsymbol{h}|}, \frac{\boldsymbol{g}}{|\boldsymbol{g}|}\right]$, via (5), up to a permutation between the two columns and up to a phase ambiguity on each column. As a result, we will be able to separate the beamforming directions from the BS to the legitimate and malicious users based on channel reciprocity.

III. Blind Channel Direction Separation

First, with our assumption on $P_{A,B}$, the columns of $[\boldsymbol{z}_1', \boldsymbol{z}_2'] \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}$ are 2×1 random vectors that range over the alphabet

$$\mathcal{Z} = \left\{\mathsf{a}\boldsymbol{z}_1' + \mathsf{b}\boldsymbol{z}_2' : \mathsf{a} \in \mathcal{A}, \mathsf{b} \in \mathcal{B}\right\}.$$

The main idea of our channel separation scheme is to use $\mathcal{Z}$ to obtain $[\boldsymbol{z}_1', \boldsymbol{z}_2']$ up to a column permutation, and then achieve the desired separation of channel directions via (5).

To explain how we may obtain $[\boldsymbol{z}_1', \boldsymbol{z}_2']$ from $\mathcal{Z}$, let us start by considering an example where both the legitimate and malicious users send BPSK symbols. That is,

$$\mathcal{Z} = \left\{\underbrace{\boldsymbol{z}_1' + \boldsymbol{z}_2'}_{\boldsymbol{v}_A}, \underbrace{-\boldsymbol{z}_1' + \boldsymbol{z}_2'}_{\boldsymbol{v}_B}, \underbrace{\boldsymbol{z}_1' - \boldsymbol{z}_2'}_{\boldsymbol{v}_C}, \underbrace{-\boldsymbol{z}_1' - \boldsymbol{z}_2'}_{\boldsymbol{v}_D}\right\}.$$

It is clear that

$$\boldsymbol{v}_A - \boldsymbol{v}_B = 2\boldsymbol{z}_1',$$
$$\boldsymbol{v}_C - \boldsymbol{v}_D = 2\boldsymbol{z}_1'.$$

This indicates that every point in $\mathcal{Z}$ is connected to another point in $\mathcal{Z}$ by a line segment along the direction of $\boldsymbol{z}_1'$. Fig. 2 denotes this observation geometrically. We will refer to this property as $\boldsymbol{z}_1'$ *covers* $\mathcal{Z}$. Similarly, it is easy to see that $\boldsymbol{v}_A - \boldsymbol{v}_C = 2\boldsymbol{z}_2'$, $\boldsymbol{v}_B - \boldsymbol{v}_D = 2\boldsymbol{z}_2'$, and hence $\boldsymbol{z}_2'$ also covers $\mathcal{Z}$. On

ISBN 978-90-827970-1-5 © EURASIP 2018

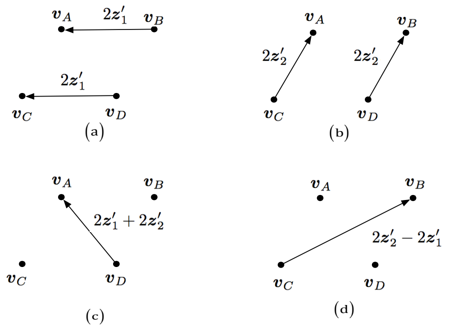

Fig. 2: Subfigures (a) and (b) show that all points of $\mathcal{Z}$ lie on lines along the directions of $\boldsymbol{z}'_1$ and $\boldsymbol{z}'_2$, respectively. Subfigures (c) and (d) show that only two points in $\mathcal{Z}$ lie on lines along the directions of $\boldsymbol{z}'_1+\boldsymbol{z}'_2$ and $\boldsymbol{z}'_2-\boldsymbol{z}'_1$, respectively.

the other hand, we have that only $\boldsymbol{v}_A - \boldsymbol{v}_D = 2\boldsymbol{z}'_1 + 2\boldsymbol{z}'_2$ and $\boldsymbol{v}_B - \boldsymbol{v}_C = 2\boldsymbol{z}'_2 - 2\boldsymbol{z}'_1$. Thus, $\boldsymbol{z}'_1+\boldsymbol{z}'_2$ covers only $\boldsymbol{v}_A$ and $\boldsymbol{v}_D$, $\boldsymbol{z}'_2-\boldsymbol{z}'_1$ covers only $\boldsymbol{v}_B$ and $\boldsymbol{v}_C$. Neither $\boldsymbol{z}'_1+\boldsymbol{z}'_2$ nor $\boldsymbol{z}'_2-\boldsymbol{z}'_1$ covers $\mathcal{Z}$. Note that the above cases exhaust the differences between all pairs of points in $\mathcal{Z}$. In summary, among all the pairwise differences, $\boldsymbol{z}'_1$ and $\boldsymbol{z}'_2$ cover $\mathcal{Z}$, but $\boldsymbol{z}'_1+\boldsymbol{z}'_2$ and $\boldsymbol{z}'_2-\boldsymbol{z}'_1$ do not. As a result, we may obtain $\boldsymbol{z}'_1$ and $\boldsymbol{z}'_2$ from $\mathcal{Z}$ by finding out all pairwise differences that cover $\mathcal{Z}$. It turns out that this observation extends to the general case as summarized in the proposition below:

Proposition 1. *Consider the general alphabet*

$$\mathcal{Z} = \{\mathsf{a}\boldsymbol{z}'_1 + \mathsf{b}\boldsymbol{z}'_2 : \mathsf{a} \in \mathcal{A}, \mathsf{b} \in \mathcal{B}\}$$

where $\mathcal{A}$ and $\mathcal{B}$ are finite and with cardinalities at least 2. Each vector in $\{\exp\{\mathrm{i}\theta_1\}\,\boldsymbol{z}'_1, \exp\{\mathrm{i}\theta_2\}\,\boldsymbol{z}'_2\}$ cover more points of $\mathcal{Z}$ than vector of the form $c_1\boldsymbol{z}'_1 + c_2\boldsymbol{z}'_2$ where c_1 and c_2 are both nonzero, θ_1 and θ_2 characterize the phase ambiguity.

Proposition 1 allows us to obtain $[\boldsymbol{z}'_1, \boldsymbol{z}'_2]$ by finding two pairwise differences of vectors in $\mathcal{Z}$ that cover most points of $\mathcal{Z}$. We will give a practical algorithm to do so later. With Proposition 1, the problem of channel separation now reduces to that of estimating the alphabet $\mathcal{Z}$ from the observation $\mathbf{Y}$.

Towards that end, further notice that the columns of $\mathbf{N}$ are i.i.d. random vectors that have the same distribution of the 2×1 random vector $\mathbf{n}$, whose elements are two independent Gaussian random variables with zero mean and variance $\frac{\sigma^2}{M}$. If the columns of $[\boldsymbol{z}'_1, \boldsymbol{z}'_2]\begin{bmatrix} \mathbf{A}\sqrt{P_1} \\ \mathbf{B}\sqrt{P_2} \end{bmatrix}$ are i.i.d. random vectors that have the same distribution as the generic 2×1 random vector $\mathbf{z}'$, then the columns of $\mathbf{Z}$, given in (4), are i.i.d. random vectors that have the same distribution as that of $\mathbf{z} = \mathbf{z}' + \mathbf{n}$. Let $F_{\mathbf{z}}$, $F_{\mathbf{z}'}$, and $F_{\mathbf{n}}$ denote the distributions of $\mathbf{z}$, $\mathbf{z}'$, and $\mathbf{n}$, respectively. Then, because $\mathbf{z}'$ and $\mathbf{n}$ are independent, we have

$$\Phi_{F_{\mathbf{z}}}(\boldsymbol{\omega}) = \Phi_{F_{\mathbf{z}'}}(\boldsymbol{\omega}) \cdot \Phi_{F_{\mathbf{n}}}(\boldsymbol{\omega}). \tag{6}$$

where $\Phi_F(\boldsymbol{\omega})$ denotes the characteristic function of the distribution F, and $\boldsymbol{\omega} = [\omega_1, \omega_2]^T$ is the 2×1 frequency vector. Note that the noise variance parameter σ^2 is a characteristic of the receiver circuitry and can be measured *a priori*. We may assume that its value is known, and thus

$$\Phi_{F_{\mathbf{n}}}(\boldsymbol{\omega}) = \exp\left\{-\frac{\sigma^2}{2M}\,|\boldsymbol{\omega}|^2\right\}$$

is also known. On the other hand, $F_{\mathbf{z}}$ can be approximated by the empirical distribution of $\mathbf{Z}$ obtained directly from the observation $\mathbf{Y}$ as in (4). Hence the distribution $F_{\mathbf{z}'}$ of $\mathbf{z}'$ can be estimated using (6).

For ease of discussion, let us use $\mathbf{z}$ to denote a generic column in the matrix $\mathbf{Z}$. To estimate $F_{\mathbf{z}'}$ efficiently, we quantize $\mathbf{z} = [z_1, z_2]^T$ and use Fast Fourier Transform (FFT) to obtain the characteristic function of the quantized version of $\mathbf{z}$ as described as follows. Consider m quantization levels $\widetilde{\mathsf{u}}_1, \widetilde{\mathsf{u}}_2, \ldots, \widetilde{\mathsf{u}}_m$, and the corresponding quantization intervals $\mathcal{B}(\widetilde{\mathsf{u}}_1), \mathcal{B}(\widetilde{\mathsf{u}}_2), \ldots, \mathcal{B}(\widetilde{\mathsf{u}}_m)$:

$$-\alpha = \widetilde{\mathsf{u}}_1 < \widetilde{\mathsf{u}}_2 < \widetilde{\mathsf{u}}_3 \ldots < \widetilde{\mathsf{u}}_{m-1} = \alpha < \widetilde{\mathsf{u}}_m,$$

$$\mathcal{B}(\widetilde{\mathsf{u}}_j) = \begin{cases} (-\infty, \widetilde{\mathsf{u}}_1], & j = 1 \\ (\widetilde{\mathsf{u}}_{j-1}, \widetilde{\mathsf{u}}_j], & j = 2, 3, \ldots, m-1 \\ (\widetilde{\mathsf{u}}_m, +\infty), & j = m \end{cases}$$

where $\widetilde{\mathsf{u}}_j - \widetilde{\mathsf{u}}_{j-1} = \triangle$ for $j = 2, 3, \ldots, m$. The elements z_1 and z_2 are respectively quantized to $\widetilde{z}_1$ and $\widetilde{z}_2$ according to:

$$\widetilde{z}_1 = \sum_{j=1}^{m} \widetilde{\mathsf{u}}_j\, \mathbb{1}(\Re\{z_1\} \in \mathcal{B}(\widetilde{\mathsf{u}}_j)) + \mathrm{i}\sum_{j=1}^{m} \widetilde{\mathsf{u}}_j\, \mathbb{1}(\Im\{z_1\} \in \mathcal{B}(\widetilde{\mathsf{u}}_j))$$

$$\widetilde{z}_2 = \sum_{j=1}^{m} \widetilde{\mathsf{u}}_j\, \mathbb{1}(\Re\{z_2\} \in \mathcal{B}(\widetilde{\mathsf{u}}_j)) + \mathrm{i}\sum_{j=1}^{m} \widetilde{\mathsf{u}}_j\, \mathbb{1}(\Im\{z_2\} \in \mathcal{B}(\widetilde{\mathsf{u}}_j))$$

where $\mathbb{1}(\cdot)$ denotes the indicator function, $\Re\{\cdot\}$ and $\Im\{\cdot\}$ takes the real part and imagery part of its input, respectively. The quantized version of $\mathbf{z}$ is then $\widetilde{\mathbf{z}} = [\widetilde{z}_1, \widetilde{z}_2]^T$. Write $\widetilde{\mathcal{U}} = \{\widetilde{\mathsf{u}}_1 + \mathrm{i}\widetilde{\mathsf{u}}_1, \widetilde{\mathsf{u}}_1 + \mathrm{i}\widetilde{\mathsf{u}}_2, \ldots, \widetilde{\mathsf{u}}_m + \mathrm{i}\widetilde{\mathsf{u}}_m\}$. Then the alphabet of $\widetilde{\mathbf{z}}$ is $\mathcal{U}^2$. We will denote it and enumerate its elements as $\widetilde{\mathcal{Z}} = \{\widetilde{\boldsymbol{u}}_1, \widetilde{\boldsymbol{u}}_2, \ldots, \widetilde{\boldsymbol{u}}_{m^4}\}$.

Let $\widetilde{\mathbf{Z}} = [\widetilde{\mathbf{z}}_1, \widetilde{\mathbf{z}}_2, \ldots, \widetilde{\mathbf{z}}_n]$, where the ith column $\widetilde{\mathbf{z}}_i$ is the quantized version of the ith column of $\mathbf{Z}$. Next, obtain the pmf of the columns of $\mathbf{Z}$ as

$$\triangle F_{\widetilde{\mathbf{z}}}(\widetilde{\boldsymbol{u}}_j) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(\widetilde{\mathbf{z}}_i = \widetilde{\boldsymbol{u}}_j) \tag{7}$$

for $j = 1, 2, \ldots, m^4$. Thus, the characteristic function of the empirical distribution $\triangle F_{\widetilde{\mathbf{z}}}$ is

$$\Phi_{\triangle F_{\widetilde{\mathbf{z}}}}(\boldsymbol{\omega}) = \sum_{j=1}^{m^4} \triangle F_{\widetilde{\mathbf{z}}}(\widetilde{\boldsymbol{u}}_j) \exp\left\{-\mathrm{i}\widetilde{\boldsymbol{u}}_j^T \boldsymbol{\omega}\right\}. \tag{8}$$

Similarly, let $\widetilde{\mathbf{Z}}' = [\widetilde{\mathbf{z}}'_1, \widetilde{\mathbf{z}}'_2, \ldots, \widetilde{\mathbf{z}}'_n]$, where the ith column $\widetilde{\mathbf{z}}'_i$ is the quantized version of the ith column of $[\boldsymbol{z}'_1, \boldsymbol{z}'_2]\begin{bmatrix} \mathbf{A}\sqrt{P_1} \\ \mathbf{B}\sqrt{P_2} \end{bmatrix}$. To do this quantization step, we employ quantization alphabet $\widetilde{\mathcal{Z}} = \{\widetilde{\boldsymbol{u}}_1, \widetilde{\boldsymbol{u}}_2, \ldots, \widetilde{\boldsymbol{u}}_{m^4}\}$. Then, the empirical pmf of the columns of $\widetilde{\mathbf{Z}}'$ is

$$\triangle F_{\widetilde{\mathbf{z}}'}(\widetilde{\boldsymbol{u}}_j) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(\widetilde{\mathbf{z}}'_i = \widetilde{\boldsymbol{u}}_j) \tag{9}$$

for $j = 1, 2, \ldots, m^4$.

Using $\Phi_{\triangle F_{\tilde{\mathbf{z}}}}(\boldsymbol{\omega})$ in place of $\Phi_{F_{\mathbf{z}}}(\boldsymbol{\omega})$ and $\triangle F_{\tilde{\mathbf{z}}'}$ in place of $F_{\mathbf{z}'}$ in (6), we obtain the estimator

$$\triangle\widehat{F}_{\tilde{\mathbf{z}}'} = \Phi^{-1}\left(\frac{\Phi_{\triangle F_{\tilde{\mathbf{z}}}}(\boldsymbol{\omega})}{\Phi_{F_{\mathrm{n}}}(\boldsymbol{\omega})}\right) \tag{10}$$

for $\triangle F_{\tilde{\mathbf{z}}'}$, where Φ^{-1} denotes the inverse Fourier Transform with respect to (8). Note that the forward and inverse Fourier Transforms in (10) can be efficiently implemented using FFT and inverse FFT, respectively.

Remark 1 The asymptotic accuracy of this estimator is guaranteed by the law of large number (LLN). Intuitively, as $n \to \infty$, $\alpha \to \infty$, $\triangle F_{\tilde{\mathbf{z}}}$ and $\triangle F_{\tilde{\mathbf{z}}'}$ respectively converges to $F_{\mathbf{z}}$ and $F_{\mathbf{z}'}$ in probability according to LLN. Notice that $\Phi(\cdot)$ and $\Phi^{-1}(\cdot)$ are orthogonal transformations, which are able to keep the convergence of its input. We would obtain that $\triangle\widehat{F}_{\tilde{\mathbf{z}}'}$ converges to $\triangle F_{\tilde{\mathbf{z}}'}$ and $F_{\mathbf{z}'}$.

On the other hand, it is observed $\mathcal{Z}$ is the support set of $F_{\mathbf{z}'}$. The convergence above-mentioned between $\triangle\widehat{F}_{\tilde{\mathbf{z}}'}$ and $F_{\mathbf{z}'}$ tells us that we may use the essential support set

$$\widehat{\mathcal{Z}} = \left\{\tilde{\boldsymbol{u}} \in \tilde{\mathcal{Z}} : \triangle\widehat{F}_{\tilde{\mathbf{z}}'}(\tilde{\boldsymbol{u}}) > \epsilon\right\}$$

of $\triangle\widehat{F}_{\tilde{\mathbf{z}}'}$ to estimate $\mathcal{Z}$. Revisiting Proposition 1, $\boldsymbol{z}_1'$ and $\boldsymbol{z}_2'$ could be obtained from $\widehat{\mathcal{Z}}$.

A. Blind channel separation algorithm

Combining Propositions 1 and (10), we can obtain the following practical channel separation algorithm:

1) Perform SVD on the observation matrix $\mathbf{Y}$, collect the first 2 left singular vectors as columns of $\boldsymbol{S}$, and obtain $\mathbf{Z} = \frac{1}{\sqrt{M}}\boldsymbol{S}^T\mathbf{Y}$.
2) Quantize the columns of $\mathbf{Z}$ to obtain $\tilde{\mathbf{Z}} = [\tilde{\mathbf{z}}_1, \tilde{\mathbf{z}}_2, \ldots, \tilde{\mathbf{z}}_n]$. Obtain the empirical pmf $\triangle F_{\tilde{\mathbf{z}}}$ of the columns of $\tilde{\mathbf{Z}}$ as described in (7).
3) Employ (10) to obtain $\triangle\widehat{F}_{\tilde{\mathbf{z}}'}$ via FFT and IFFT.
4) Choose a small $\epsilon > 0$. Obtain the essential support set $\widehat{\mathcal{Z}} = \left\{\tilde{\boldsymbol{u}} \in \tilde{\mathcal{Z}} : \triangle\widehat{F}_{\tilde{\mathbf{z}}'}(\tilde{\boldsymbol{u}}) > \epsilon\right\}$. Write $\widehat{\mathcal{Z}} = \left\{\widehat{\boldsymbol{v}}_1, \widehat{\boldsymbol{v}}_2, \ldots, \widehat{\boldsymbol{v}}_{|\widehat{\mathcal{Z}}|}\right\}$.
5) Obtain the set of pairwise differences $\mathcal{D} = \left\{\widehat{\boldsymbol{v}}_i - \widehat{\boldsymbol{v}}_j : i < j \text{ and } i, j \in \{1, \ldots, |\widehat{\mathcal{Z}}|\}\right\}$. Choose a small $\gamma > 0$. Find a subset $\mathcal{D}^* = \{\boldsymbol{d}_1, \boldsymbol{d}_2, \ldots, \boldsymbol{d}_T\} \subseteq \mathcal{D}$ that partitions $\mathcal{D}$ into T parts, i.e, $\mathcal{D} = \bigcup_{t=1}^{T}\mathcal{D}_t$ and $\mathcal{D}_s \cap \mathcal{D}_t = \emptyset$ for any $s \neq t$, such that $\mathcal{D}_t = \left\{\boldsymbol{d} \in \mathcal{D} : \left|\frac{\boldsymbol{d}^H\boldsymbol{d}_t}{|\boldsymbol{d}||\boldsymbol{d}_t|} - 1\right| \leq \gamma\right\}$ for $t = 1, 2\ldots, T$. Note that T is the minimum number of parts ($\mathcal{D}_t$'s) that are needed to partition $\mathcal{D}$. Its value depends on the choice of γ.
6) For each $t \in \{1, 2, \ldots, T\}$, if $\widehat{\boldsymbol{v}}_i - \widehat{\boldsymbol{v}}_j \in \mathcal{D}_t$, then we collect $\widehat{\boldsymbol{v}}_i$ and $\widehat{\boldsymbol{v}}_j$ in $\widehat{\mathcal{Z}}_t$. Define the weight of $\boldsymbol{d}_t$ as $W(\boldsymbol{d}_t) = \sum_{\widehat{\boldsymbol{v}} \in \widehat{\mathcal{Z}}_t} \triangle\widehat{F}_{\tilde{\mathbf{z}}'}(\widehat{\boldsymbol{v}})$. Use the 2 vectors in $\mathcal{D}^*$ with the largest weights as estimates of the columns of $[\boldsymbol{z}_1', \boldsymbol{z}_2']$.
7) Employ (5) to estimate $\left[\frac{\boldsymbol{h}}{|\boldsymbol{h}|}, \frac{\boldsymbol{g}}{|\boldsymbol{g}|}\right]$, up to a permutation of the columns and up to a phase ambiguity on each column, from the estimated $[\boldsymbol{z}_1', \boldsymbol{z}_2']$ obtained in step 6).

Note that steps 5) and 6) effectively find all the pairwise difference vectors of $\mathcal{Z}$ that cover most points of $\mathcal{Z}$. As we employ $\widehat{\mathcal{Z}}$ to approximate $\mathcal{Z}$, $\widehat{\mathcal{Z}}$ often contains many more points than $\mathcal{Z}$. To solve this problem, step 5) first clusters all pairwise differences. Step 6) uses a likelihood metric to approximately obtain the set of covering pairwise difference vectors among the clusters obtained in step 5).

Remark 2 The proposed algorithm obtains partial CSI. Comparing with the full CSI, i.e., $[\boldsymbol{h}, \boldsymbol{g}]$, the obtained result has permutation of the columns and phase ambiguity on each column. This is similar to some blind source separation methods (e.g., ICA), and needs further steps to obtain full CSI. Nevertheless, unlike existing methods, it is clear that the proposed algorithm requires no *a priori* knowledge of $\mathcal{Z}$ or the symbol alphabets of the legitimate and malicious users, and imposes no restrict on statistic dependence between symbols of the legitimate and malicious users.

IV. Performance Evaluation

In this section, we present some simulation results to evaluate the performance of the proposed blind channel separation (BCS) scheme. We start by describing the metric that we employ to measure the channel seperation performance.

Suppose that $\boldsymbol{d}$ is an uplink channel direction vector obtained by some channel estimation algorithm. By reciprocity, downlink beamforming is performed based on $\boldsymbol{d}$. Then the leakage-to-direct power ratio (LDPR) $\frac{|\frac{\boldsymbol{g}^T}{|\boldsymbol{g}|}\boldsymbol{d}|^2}{|\frac{\boldsymbol{h}^T}{|\boldsymbol{h}|}\boldsymbol{d}|^2}$ measures the ratio of the power leaked to the malicious user to the power directed towards the legitimate user if $\boldsymbol{d}$ is employed to perform beamforming. Clearly, $\frac{|\frac{\boldsymbol{g}^T}{|\boldsymbol{g}|}\frac{\boldsymbol{h}}{|\boldsymbol{h}|}|^2}{|\frac{\boldsymbol{h}^T}{|\boldsymbol{h}|}\frac{\boldsymbol{h}}{|\boldsymbol{h}|}|^2}$ is the LDPR value when the channel direction estimation is perfect.

Now, let $[\widehat{\boldsymbol{h}}_1, \widehat{\boldsymbol{h}}_2]$ be the channel direction vectors estimated using the blind channel separation scheme described in Section III. Since we do not know whether $\widehat{\boldsymbol{h}}_1$ or $\widehat{\boldsymbol{h}}_2$ corresponds to $\boldsymbol{h}$, we consider the minimum between the LDPRs of the two possibilities [1] :

$$\mathrm{LDPR}(\boldsymbol{h}, \boldsymbol{g}) \triangleq \min\left\{\frac{|\frac{\boldsymbol{g}^T}{|\boldsymbol{g}|}\widehat{\boldsymbol{h}}_1|^2}{|\frac{\boldsymbol{h}^T}{|\boldsymbol{h}|}\widehat{\boldsymbol{h}}_1|^2}, \frac{|\frac{\boldsymbol{g}^T}{|\boldsymbol{g}|}\widehat{\boldsymbol{h}}_2|^2}{|\frac{\boldsymbol{h}^T}{|\boldsymbol{h}|}\widehat{\boldsymbol{h}}_2|^2}\right\}. \tag{11}$$

Note that the LDPR in (11) is a function of the channel vectors $[\boldsymbol{h}, \boldsymbol{g}]$. In the simulation, we generate $10,000$ instances of the channel vectors based on the block Rayleigh fading model. That is, the elements of $\boldsymbol{h}$ and $\boldsymbol{g}$ are chosen as i.i.d. CSCG random variables with 0 mean and variance 1. We then average the LDPRs given by (11) over the $10,000$ channel instances.

In the simulation, we set $P_1 = P_2$. The legitimate user sends equally likely random BPSK symbols, i.e., $P_A(-1) =$

[1]The channels could be identified by some *asymmetry* measures, which is beyond the scope of this paper due to space limitation. We indeed assume perfect channel identification herein.

ISBN 978-90-827970-1-5 © EURASIP 2018

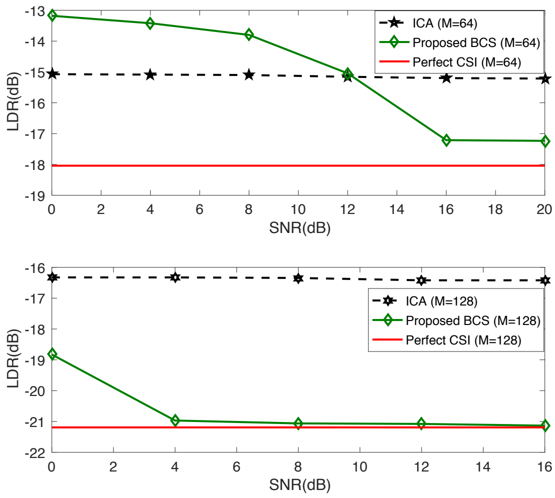

Fig. 3: Average LDPR obtained by the proposed blind channel separation algorithm in an MMIMO system respectively with 64 and 128 antennas at the BS.

$P_A(+1) = \frac{1}{2}$. The malicious user sends BPSK symbols according to $P_{B|A}(+1|+1) = P_{B|A}(-1|+1) = \frac{1}{2}$, $P_{B|A}(+1|-1) = \frac{1}{3}$, $P_{B|A}(-1|-1) = \frac{2}{3}$. The case of 500 random BPSK symbols ($n = 500$) is simulated. The proposed BCS scheme is performed according to Section III.A.

Fig. 3 plots the average LDPRs obtained by the proposed blind channel separation algorithm for the case of $n = 500$ versus the SNR in an MMIMO system with $M = 64$ and $M = 128$ antennas at the BS, respectively. Also plotted in the figure are the LDPR values obtained by ICA scheme and by using perfect CSI information, respectively.

For the MMIMO system with $M = 64$, we observe from the first subfigure of Fig. 3 that the proposed algorithm achieves LDPR close to perfect LDPR performance within 1dB when the per-antenna SNR reaches 12 dB. In contrast, the LDPR achieved by traditional ICA scheme is 3dB worse than that achieved by perfect CSI. For $M = 128$, we observe from the second subfigure of Fig. 3 that the proposed algorithm provides almost perfect LDPR performance when the per-antenna SNR reaches 4dB. Meanwhile, it outperforms ICA scheme up to 5dB. It indicates the proposed scheme is effective to separate channel directions when the malicious and legitimate symbols are dependent, and its distributions are unknown to the BS. Due to space limitation, we cannot present more results.

V. Conclusions

We have proposed a blind channel direction separation algorithm to differentiate the channel directions from a legitimate user and a malicious user to the BS in the uplink of a massive MIMO system. With channel reciprocity, the BS then will be able to use the channel directions to beamform to the legitimate and malicious users separately and further verify their identities by use of higher layer authentication protocols. Extensions of the proposed blind channel separation to the cases of multiple legitimate and malicious users are of interest.

References

[1] K. N. R. S. V. Prasad, E. Hossain and V. K. Bhargava, "Energy Efficiency in Massive MIMO-Based 5G Networks: Opportunities and Challenges," in *IEEE Wireless Communications*, vol.PP, no.99, pp.2-10, 2017.

[2] M. Shafi et al., "5G: A Tutorial Overview of Standards, Trials, Challenges, Deployment, and Practice," in *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1201-1221, June 2017.

[3] Y. O. Basciftci, C. E. Koksal and A. Ashikhmin, "Securing massive MIMO at the physical layer," in *2015 IEEE Conference on Communications and Network Security (CNS)*, Florence, 2015, pp. 272-280.

[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," in *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 5, pp. 932-940, May 2015.

[5] D. Kapetanovic, A. Al-Nahari, A. Stojanovic and F. Rusek, "Detection of active eavesdroppers in massive MIMO," *2014 IEEE 25th Annual International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC)*, Washington DC, June 2014, pp. 585-589.

[6] J. M. Kang, C. In and H. M. Kim, "Detection of Pilot Contamination Attack for Multi-Antenna Based Secrecy Systems," *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, Glasgow, 2015, pp. 1-5.

[7] J. K. Tugnait, "Detection of pilot contamination attack in T.D.D./S.D.M.A. systems," *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, May 2016, pp. 3576-3580.

[8] J. K. Tugnait, "On detection of pilot contamination attack in multiple antenna systems," *2015 49th Asilomar Conference on Signals, Systems and Computers,* Pacific Grove, CA, May 2015, pp. 1757-1761.

[9] J. K. Tugnait, "Self-Contamination for Detection of Pilot Contamination Attack in Multiple Antenna Systems," in *IEEE Wireless Communications Letters*, vol. 4, no. 5, pp. 525-528, Oct. 2015.

[10] D. Kapetanovi, G. Zheng, K. K. Wong and B. Ottersten, "Detection of pilot contamination attack using random training and massive MIMO," *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, London, 2013, pp. 13-18.

[11] J. Vinogradova, E. Björnson and E. G. Larsson, "Detection and mitigation of jamming attacks in massive MIMO systems using random matrix theory," *2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Edinburgh, 2016, pp. 1-5.

[12] X. Tian, M. Li and Q. Liu, "Random-Training-Assisted Pilot Spoofing Detection and Security Enhancement," in IEEE Access, vol. 5, pp. 27384-27399, 2017.

[13] J. Park, S. Yun and J. Ha, "Detection of pilot contamination attack in the MU-MISOME broadcast channels," *2016 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju, 2016, pp. 664-666.

[14] J. Yang, S. Xie, X. Zhou, R. Yu and Y. Zhang, "A Semiblind Two-Way Training Method for Discriminatory Channel Estimation in MIMO Systems," in *IEEE Transactions on Communications*, vol. 62, no. 7, pp. 2400-2410, July 2014.

[15] Q. Xiong, Y. C. Liang, K. H. Li, Y. Gong and S. Han, "Secure Transmission Against Pilot Spoofing Attack: A Two-Way Training-Based Scheme," in *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 5, pp. 1017-1026, May 2016.

[16] F. Bai, P. Ren, Q. Du and L. Sun, "A hybrid channel estimation strategy against pilot spoofing attack in MISO system," *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Valencia, 2016, pp. 1-6.

[17] D. Xu, P. Ren, Y. Wang, Q. Du and L. Sun, "ICA-SBDC: A channel estimation and identification mechanism for MISO-OFDM systems under pilot spoofing attack," in *2017 IEEE International Conference on Communications (ICC)*, Paris, 2017, pp. 1-6.

[18] J. K. Tugnait, "On Detection and Mitigation of Reused Pilots in Massive MIMO Systems," in *IEEE Transactions on Communications*, Early Access, 2018.

[19] R. R. Müller, L. Cottatellucci and M. Vehkaperä, "Blind Pilot Decontamination," in *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 773-786, Oct. 2014.

ISBN 978-90-827970-1-5 © EURASIP 2018