IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 27, 2019, accepted October 12, 2019, date of publication October 17, 2019, date of current version October 30, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2948078

Proactive Eavesdropping via Covert Pilot Spoofing Attack in Multi-Antenna Systems

XINGBO LU, WEIWEI YANG, (Member, IEEE), YUEMING CAI, (Senior Member, IEEE), AND XINRONG GUAN

Institute of Communications Engineering, Army Engineering University of PLA, Nanjing 210007, China

Corresponding author: Weiwei Yang (wwyang1981@163.com)

This work was supported by the National Natural Science Foundation of China under Grant 61771487 and Grant 61471393.

ABSTRACT Proactive eavesdropping is an effective method for government to monitor suspicious users who are deemed to misuse communication systems for illegal activities. In this paper, considering that a legitimate full-duplex (FD) eavesdropper tries to monitor a dubious multi-antenna system, we propose a covert pilot spoofing attack (PSA) scheme to enhance the legitimate eavesdropping performance by taking the channel training phase into consideration. For the proposed covert PSA scheme, the total error detection probability and optimal detection threshold of suspicious source are derived as the worst case for the considered monitoring system. Given the optimal detection threshold, the closed-form expressions of effective eavesdropping rate are also derived based on the results of detection at suspicious source. Furthermore, an optimal power allocation algorithm to maximize the effective eavesdropping rate is proposed under the covert PSA and transmission power constraints. Simulation results illustrate that the adversary's uncertain about channel state information (CSI) before channel estimation process, can be exploited by legitimate eavesdropper to cover the PSA without being detected. Therefore, the proposed covert PSA scheme can achieve a better performance with respect to effective eavesdropping rate and effectively combat with a suspicious multi-antenna system.

INDEX TERMS Proactive eavesdropping, covert communication, pilot spoofing attack, channel uncertainty.

I. INTRODUCTION

With the development of wireless communication technologies, various wireless communication applications and services profoundly change our daily life [1], [2]. On the other hand, private information security has attracted utmost concern due to the intrinsic open nature of wireless air interfaces. In response to this, physical layer security (PLS), which utilizes signal processing techniques and wireless fading channels to safeguard private information confidentiality, has also attracted much research attentions [3]–[13]. With respect to PLS, a great number of signal processing techniques, including artificial noise (AN) [3]–[5], oriented beamforming [6], [7] and cooperative jamming [8], have been exploited to enhance security performance in the wireless communication systems [9]–[13].

It is noted that most existing works focus on preventing the private information leaking to eavesdroppers, however, such ubiquitous communication systems and sophisticated signal processing techniques may be illegally used by criminals or terrorists, which severely jeopardize the public safety [14]. For example, terrorists may exchange the suspicious information to facilitate hijacking through core infrastructures that owned by themselves, where the suspicious source equips with multi-antennas and sophisticated PLS techniques are used. In viewing of the security threat posed by suspicious communication systems, in this paper, we investigate how to eavesdrop the suspicious messages in a multi-antenna system through legitimate monitoring, which is a paradigm shift in PLS from the protection of authorised communications to the surveillance of suspicious communications.

A. RELATED WORKS AND MOTIVATIONS

A straightforward way to legitimately monitor suspicious communication is passive eavesdropping, where the eavesdropper only silently overheads the wireless channels. However, the eavesdropper may be located far away from the suspicious transmitter and the wiretap channel is worse than suspicious channel, which makes the legitimate

The associate editor coordinating the review of this manuscript and approving it for publication was Khaled Rabie.

eavesdropping more challenge. To enhance the eavesdropping performance, a proactive eavesdropping paradigm was first proposed in [15], where a full-duplex (FD) eavesdropper tried to monitor and intervened a pair of suspicious users via active jamming. With a assumption that an adaptive rate strategy is adopted for suspicious source to send suspicious messages, the legitimate eavesdropper can successfully overhear the intercepted information only when the wiretap channel capacity is lager than the suspicious communication rate. And the effective eavesdropping rate, which defined as suspicious communication rate that satisfies such a condition, was coined to evaluate the eavesdropping performance. For the same system, a robust proactive eavesdropping scheme against imperfect self-interference cancelation and a two-player noncooperative game power allocation approach were proposed in [16] and [17], respectively. In [18], considering that a multi-antenna FD eavesdropper was a spoofing relay, its power allocation and beamforming vector were jointly optimized to maximize the eavesdropping rate. Related works [19]–[24] extended the proactive eavesdropping paradigm to wireless powered communication network, unmanned aerial vehicles and suspicious relay systems. All the aforementioned works, however, have assumed that the suspicious entities equipped with single antenna, which limits the suspicious nodes to combat with proactive eavesdropper. Nevertheless, the legitimate eavesdropping is very challenging when the suspicious source is equipped with multi-antenna, since it offers spatial degrees of freedom. In order to successfully eavesdrop a multi-antenna system, the optimal jamming power and beamforming vectors were jointly designed for eavesdropping in [25]. However, it will be invalid with the antennas of suspicious source increasing since little private information leaks to eavesdropper through security-oriented beamforming [26]. So, the effective eavesdropping rate may be still zero even the eavesdropper jams with the maximal jamming power.

Noted that above works focus on improving the legitimate eavesdropping performance during the data transmission phase. In practice, the legitimate monitoring performance can be strengthened during channel training phase, especially in multi-antenna systems, since the security-oriented beamforming in multi-antenna systems highly depends on the uplink channel training process. This motivates the legitimate eavesdropper to launch active attacks (e.g. pilot attack [27], [28] and jamming attack [29]) during the training phase for contaminating channel training and altering the downlink beam pattern. As a result, the beamforming based on a weighted channel state information (CSI) of suspicious channel and wiretap channel will be directed to both the suspicious destination and legitimate eavesdropper [27], [28]. And a pilot spoofing attack (PSA) scheme has been proposed in [29] to enhance monitoring performance. However, the PSA can be detected with a high probability through energy-ratio-based detector since the strategy of PSA is exposed to suspicious source [30]–[32]. More importantly, the suspicious source could estimate both suspicious and wiretap channels if the PSA is correctly detected. Furthermore, by exploiting beamforming and artificial noise, the data can be securely transmitted once the wiretap channel is exposed to suspicious source. Hence, to successfully monitor a suspicious multi-antenna system, the PSA needs to be tactfully hidden without being detected.

Covert communication, which utilizes the various uncertainties to disturb warden, can hide the existence of communication from warden and achieve a positive covert rate. In [33], an uninformed jammer was hired to assist Alice and Bob. With the help of uninformed jammer, Alice can communicate covertly with Bob in the presence of a watchful adversary Willie. In [37], a full-duplex receiver was used to achieved covert wireless communication. In [38], the confusion introduced by uncertainty of channel was exploited to achieve covert communication in relay network. It has been proven that the uncertainties in terms of jamming power [33]–[37], receiver noise power [39], [40] and the wireless fading channel [38], [41] can be exploited to achieve a positive covert communication rate, which provides a promising approach to launch active attacks without being detected for proactive eavesdropper. Since the suspicious source, who plays the part of warden to detect whether there is a PSA during the training phase, is uncertain even completely unknown about CSI, this gives a chance for eavesdropper to launch PSA without being detected. **However, whether the covert PSA is achievable and how much pilot power can be covered in the uncertain of channel are undefined. Furthermore, how much the legitimate monitoring performance can be improved by introducing covert PSA remains to be seen.** To our best knowledge, few works has focus on addressing these questions, which motivates us to design a covert PSA scheme for further improving monitoring performance in a suspicious multi-antenna system.

B. OUR APPROACH AND CONTRIBUTIONS

Motivated by the above analysis, in this paper, we propose a covert PSA scheme to assist proactive eavesdropping, in which the PSA can be covered without being detected by exploiting the uncertainty of channels during the uplink training phase. In the considered suspicious communication scenario, where the suspicious source is equipped with multi-antennas, we also make use of a FD eavesdropper to achieve proactive eavesdropping. Specifically, the legitimate eavesdropper is a pretender to send pilot symbols as that of suspicious destination during the training phase. As the practicable covert PSA power may be less than the pilot power of suspicious destination and most beams still be directed to suspicious destination, a active jamming scheme is also adopted to improve the proactive eavesdropping performance during data transmission phase.

The main contributions of this work are listed as follows.

- For the first time, we propose a covert PSA scheme to enhance the legitimate eavesdropping performance in a multi-antenna system by taking the uncertainty of

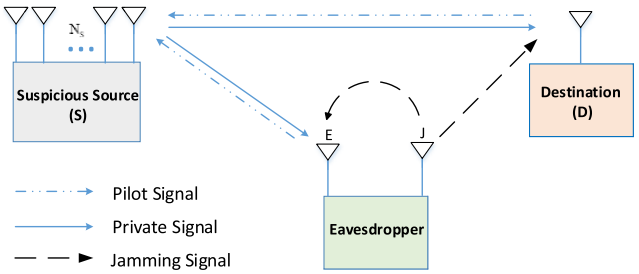

FIGURE 1. System model.

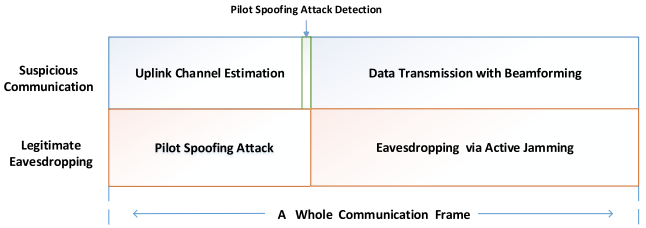

FIGURE 2. A whole transmission time slot.

channels at suspicious source into consideration during the uplink training phase.

- We analyze the total error detection performance and derive the optimal detection threshold for suspicious source. We also derive the closed-form expression of effective eavesdropping rate with the proposed covert PSA scheme. An optimal power allocation algorithm to maximize the effective eavesdropping rate is proposed with the covert PSA and transmission power requirement.
- Numerical results demonstrate that the uncertainty of channels during the training phase can be utilized to achieve covert PSA for legitimate eavesdropper. Furthermore, by comparing to the proactive eavesdropping scheme without PSA, the proposed scheme can achieve a better performance with respect to effective eavesdropping rate and can effectively combat with multi-antenna suspicious communication system.

The rest of the this paper is organized as follows: Section II describes the system model. In section III, the proposed covert PSA scheme and performance of PSA detection at suspicious source are analyzed. Effective eavesdropping rate analysis and optimization under two different scenarios in data transmission phase are given in section III. Simulation results are provided in section V. Finally, we draw conclusions in section VI.

II. SYSTEM MODEL

We consider a proactive eavesdropping paradigm as shown in Fig.1, in which a suspicious source (S) equipped with N_S antennas sends suspicious messages to a single antenna destination (D). A legitimate FD eavesdropper equipped with two antennas, one for eavesdropping (E) and another for jamming (J), tries to eavesdrop the suspicious messages. We assume that all channels are mutually independent and follow quasi-static Rayleigh fading, which indicates that the channel coefficients are invariant within a time slot, but independently change from one frame to another. Specifically, the $S \rightarrow D$ and $S \rightarrow E$ channels are respectively modeled as narrowband 2-D spatial model. And it can be expressed as $\mathbf{h}_{SI} = \sqrt{\frac{N_S}{\beta_{SI} L_{SI}}} \sum_{l=1}^{L_{SI}} \alpha_l \mathbf{a}(\theta_l)$, where $I \in \{E, D\}$, β_{SI} denotes large-scale fading, L_{SI} is the number of multipath, α_l denotes small-scale fading of lth path that assumed to be complex Gaussian random variable with zero-mean and unit variance and $\mathbf{a}(\theta_l)$ represents array steering vector. For an uniform linear array, $\mathbf{a}(\theta_l) = \frac{1}{\sqrt{N_S}} \left[1, e^{j\frac{2\pi}{\lambda} d \sin(\theta_l)}, \cdots, e^{j\frac{2\pi}{\lambda}(N_S-1)d \sin(\theta_l)}\right]^T$ [42]. The interference channels $J \rightarrow D$ and $J \rightarrow E$ are denoted by h_{JD} and h_{JE}, where the mean of $|h_{JD}|^2$ and $|h_{JE}|^2$ are $\frac{1}{\beta_{JD}}$ and $\frac{1}{\beta_{JE}}$, respectively. It's worth to note that the self-interference (SI) channel h_{JE} is also modeled via Rayleigh distribution since the antenna E and J are isolated and the major interference comes from scattering. Furthermore, perfect SI cancelation is difficult due to the finite resolution of analog-to-digital converter, however, partial SI can be cancelled in such a FD legitimate eavesdropper.

We assume that time division duplex (TDD) protocol is adopted and each time slot is divided into two phases, including channel training phase and data transmission phase, as shown in Fig.2. During channel training phase, D broadcasts common pilot symbols to S for estimating the channel. While, the legitimate eavesdropper also broadcasts synchronized and identical pilot symbols as that of D through antenna E to contaminate channel estimation, but antenna J keeps silence to reduce the chance of exposure. The suspicious source S, who plays the role of a warden during channel training phase, tries to estimate channel $\mathbf{h}_{SD}$ and detect any PSA or jamming attack based on the received pilot symbols. We also assume that the distribution of channels is known to S, but the exact CSI is unaware during the training phase, since the distribution of channels can be acquired through long-time statistic, while exact CSI need to be instantaneously estimated. In addition, we assume that S has knowledge of the transmission power and noise variance.

After receiving the pilot symbols, S needs to make a decision as to whether a PSA happens. Then, according to the results of channel estimation and active attack detection, S transmits data with beamforming during data transmission phase. Specifically, we assume that a maximum-ratio-transmission (MRT) beamforming is adopted when S deems that there is absence of PSA, but the oriented beamforming, where the suspicious messages and interference signals are respectively directed to D and E, is adopted when the PSA is exposed to S. To effectively improve the eavesdropping performance, the legitimate eavesdropper sends interference signals through the antenna J to decrease the achievable

data rate at D, while decoding the suspicious communication through the antenna E during data transmission phase. In the following, we will introduce the proposed proactive eavesdropping scheme under a multi-antennas system.

III. PROACTIVE EAVESDROPPING VIA COVERT PSA

A. PSA DETECTION AT SUSPICIOUS SOURCE

As described in section II, after receiving the pilot symbols, S needs to make a decision as to whether a PSA happens, which is a binary hypothesis testing problem. For the simplification of expression, two hypothesises are defined, where H_0 represents that there is absence of PSA and H_1 means there is existing PSA in channel training phase. We define the detection error of PSA in the aspect of false alarm probability (p_{fa}) and miss detection probability (p_{md}). Specifically, p_{fa} is the probability that S agrees H_1, while H_0 is true, oppositely, p_{md} is the probability that S agrees H_0, while H_1 is true. Let p_0 and p_1 denote the priori probabilities of H_0 and H_1, respectively. The total probability of detection error can be defined as

$$\xi = p_0 p_{fa} + p_1 p_{md} \tag{1}$$

To achieve a covert PSA, the total probability of detection error should be satisfied that $\xi = p_0 p_{fa} + p_1 p_{md} \geq \min\{p_0, p_1\} - \varepsilon$ even under the optimal detection threshold, for any $\varepsilon > 0$ [38].

For effectively cover the PSA, we derive the optimal detection threshold of S to minimize the total probability of detection error, which is a worst case for proactive eavesdropper. Specifically, under the hypothesise H_0 and H_1, the collected pilot symbol at S during the channel training phase can be respectively expressed as

$$H_0: \ \mathbf{y}_0 = \sqrt{p_D}\mathbf{h}_{SD}x + \mathbf{n} \tag{2}$$

$$H_1: \ \mathbf{y}_1 = \sqrt{p_D}\mathbf{h}_{SD}x + \sqrt{p_E}\mathbf{h}_{SE}x + \mathbf{n} \tag{3}$$

where x is the pilot symbol that satisfies $E\left\{x^H x\right\} = 1$, P_D and P_E respectively denote the transmission power of pilot symbols at D and E, and $\mathbf{n} \in \mathbb{C}^{N_S \times 1}$ is the noise vector that its elements are independent identically distributed (i.i.d) and satisfied $\mathcal{CN}\left(0, \sigma_0^2\right)$. During the channel training phase, we assume the instantaneous channels are uncertainty but the complex statistic CSI under both hypotheses are certain for S. To detect whether a PSA happens, a optimal detection scheme, i.e. radiometer [37], is adopted at S to distinguish H_0 and H_1, and a test statistic T is defined as

$$T_i(m) = \frac{1}{m}\sum_{w=1}^{m}\left\|\mathbf{y}_i^{(w)}\right\|^2 \underset{D_1}{\overset{D_0}{\lesseqgtr}} \gamma, \quad (i = 0, 1) \tag{4}$$

where m is the number of the samples, D_1 and D_0 are binary decisions of S that inter there is a PSA or not, and γ is the test threshold.

Without loss of generality, we assume that $p_0 = p_1 = 0.5$, which means S has no priori knowledge about PSA strategy of E. To achieve covert communication, for any $\varepsilon > 0$, the total probability of detection error $\xi = p_{fa} + p_{md} \geq 1 - \varepsilon$ needs to be ensured when $m \to \infty$. And the probability of false alarm under H_0 is

$$\begin{aligned} p_{fa} &= \Pr\{T_0(m) > \gamma \mid H_0\} \\ &= \Pr\left\{\left(\sigma_1^2 + \sigma_0^2\right)\frac{\chi^2_{2mN_S}}{m} > \gamma\right\} \end{aligned} \tag{5}$$

where $\sigma_1^2 = p_D|h_{SD}|^2$ is the received power of pilot symbol for one of antennas at S under H_0 and $\chi^2_{2mN_S}$ denotes chi-squared distribution with the freedom $2mN_S$. According to the Strong Law of Large Numbers [32-34], we have

$$\Pr\left\{\lim_{m\to\infty}\frac{\chi^2_{2mN_S}}{m} = N_S\right\} = 1 \tag{6}$$

By considering $m \to \infty$ and substituting (6) into (5), the probability of false alarm under H_0 can be computed as

$$\begin{aligned} p_{fa} &= \Pr\left\{\sigma_1^2 > \frac{\gamma}{N_S} - \sigma_0^2\right\} \\ &= \begin{cases} 1, & \gamma < N_S\sigma_0^2 \\ \exp\left(-\dfrac{\gamma - N_S\sigma_0^2}{N_S\bar{\sigma}_1^2}\right), & \gamma \geq N_S\sigma_0^2 \end{cases} \end{aligned} \tag{7}$$

where $\bar{\sigma}_1^2 = \frac{p_D}{\beta_{SD}}$. Similarly, we can derive the probability of missed detection under H_1 as

$$\begin{aligned} p_{md} &= \Pr\{T_1(m) < \gamma \mid H_1\} \\ &= \Pr\left\{\left(\sigma_1^2 + \sigma_2^2 + \sigma_0^2\right)\frac{\chi^2_{2mN_S}}{m} < \gamma\right\} \\ &= \begin{cases} 0, & \gamma < N_S\sigma_0^2 \\ \Pr\left\{\theta_1 < \dfrac{\gamma}{N_S} - \sigma_0^2\right\}, & \gamma \geq N_S\sigma_0^2 \end{cases} \end{aligned} \tag{8}$$

where $\theta_1 = \sigma_1^2 + \sigma_2^2 = p_D|h_{SD}|^2 + p_E|h_{SE}|^2$ is the power of received pilot symbol for one of antennas at S under H_1. And it follows generalized chi-squared distribution with the probability density function as [38]

$$f\left(x, K, \bar{\sigma}_1^2, \bar{\sigma}_2^2, \cdots, \bar{\sigma}_K^2\right) = \sum_{k=1}^{K} c_k \exp\left(-\frac{x}{\bar{\sigma}_k^2}\right) \tag{9}$$

where

$$c_k = \frac{1}{\bar{\sigma}_k^2 \prod_{l=1, l\neq k}^{K}\left(1 - \bar{\sigma}_l^2/\bar{\sigma}_k^2\right)} \tag{10}$$

For the considered system, we have $k = 2$, $\bar{\sigma}_2^2 = \frac{p_E}{\beta_{SE}}$, $c_1 = \frac{\beta_{SE}\beta_{SD}}{\beta_{SE}p_D - \beta_{SD}p_E}$, $c_2 = \frac{\beta_{SE}\beta_{SD}}{\beta_{SD}p_E - \beta_{SE}p_D}$ and p_{md} can be computed as

$$p_{md} = \begin{cases} 0, & \gamma < N_S\sigma_0^2 \\ \displaystyle\sum_{k=1}^{2}\left(c_k\bar{\sigma}_k^2 - c_k\bar{\sigma}_k^2\exp\left(-\frac{\gamma - N_S\sigma_0^2}{N_S\bar{\sigma}_k^2}\right)\right), & \gamma \geq N_S\sigma_0^2 \end{cases} \tag{11}$$

As p_{fa} and p_{md} in (8) and (11) are functions of γ and $\gamma > 0$, we assume that S can set the optimal threshold γ^* to minimize the total error detection probability. Specifically, when

$\gamma < N_S\sigma_0^2, \xi = 1$ is permanent as $p_{fa} = 1$. When $\gamma \geq N_S\sigma_0^2$, to examine a optimal threshold γ^* at S, we derive the first derivative of ξ as

$$\begin{aligned}\frac{\partial \xi}{\partial \gamma} &= \frac{\partial\left(\exp\left(-\frac{\gamma-N_S\sigma_0^2}{N_S\bar{\sigma}_1^2}\right)+\sum_{k=1}^{2}\left(c_k\bar{\sigma}_k^2-c_k\bar{\sigma}_k^2\exp\left(-\frac{\gamma-N_S\sigma_0^2}{N_S\bar{\sigma}_k^2}\right)\right)\right)}{\partial\gamma} \\ &= \sum_{k=1}^{2}\left(\frac{c_k}{N_S}\exp\left(-\frac{\gamma-N_S\sigma_0^2}{N_S\bar{\sigma}_k^2}\right)\right)-\frac{1}{N_S\bar{\sigma}_1^2}\exp\left(-\frac{\gamma-N_S\sigma_0^2}{N_S\bar{\sigma}_1^2}\right) \\ &\overset{(a)}{=} \left(\frac{c_1\bar{\sigma}_1^2-1}{N_S\bar{\sigma}_1^2}\right)\exp\left(-\frac{\gamma-N_S\sigma_0^2}{N_S\bar{\sigma}_1^2}\right)-\frac{c_1}{N_S}\exp\left(-\frac{\gamma-N_S\sigma_0^2}{N_S\bar{\sigma}_2^2}\right)\end{aligned} \tag{12}$$

where (a) is because $c_2 = -c_1$. Let $\frac{\partial\xi}{\partial\gamma} = 0$, then we have

$$\bar{\gamma} = \frac{N_S\bar{\sigma}_1^2\bar{\sigma}_2^2}{\bar{\sigma}_2^2-\bar{\sigma}_1^2}\ln\left(\frac{c_1\bar{\sigma}_1^2-1}{c_1\bar{\sigma}_1^2}\right)+N_S\sigma_0^2 \tag{13}$$

Notice that $\frac{\partial\xi}{\partial\gamma} \leq 0$ for $N_S\sigma_0^2 < \gamma \leq \bar{\gamma}$ and $\frac{\partial\xi}{\partial\gamma} \geq 0$ for $\bar{\gamma} \leq \gamma$. So the optimal threshold when $\bar{\sigma}_1^2 \neq \bar{\sigma}_2^2$ is

$$\gamma^* = \frac{N_S\bar{\sigma}_1^2\bar{\sigma}_2^2}{\bar{\sigma}_1^2-\bar{\sigma}_2^2}\ln\left(\frac{c_1\bar{\sigma}_1^2-1}{c_1\bar{\sigma}_1^2}\right)+N_S\sigma_0^2 \tag{14}$$

Note that for $\bar{\sigma}_1^2 = \bar{\sigma}_2^2$, we have

$$\frac{\partial\xi}{\partial\gamma} = \frac{-1}{N_S\bar{\sigma}_1^2}\exp\left(-\frac{\gamma-N_S\sigma_0^2}{N_S\bar{\sigma}_1^2}\right) \tag{15}$$

Since $\frac{\partial\xi}{\partial\gamma} < 0$ in (15), so the optimal detection threshold is $\gamma \rightarrow +\infty$. According to the received pilot symbols and optimal detection threshold, S can decide whether a PSA happens.

Remark 1: According to the false alarm probability p_{fa}, missed detection probability p_{md} and the optimal threshold γ^*, the total probability of detection error ξ has nothing to do with the number of antennas N_S. In addition, the false alarm probability p_{fa} is irrelevant to the power of PSA.

B. PROACTIVE EAVESDROPPING DURING DATA TRANSMISSION PHASE

It is clear that S may adopt different beamforming scheme to transmit suspicious messages during data transmission phase, since different CSI is acquired according to the detection result. We assume that a MRT beamforming is adopted when S deems that there is absence of PSA, as MRT beamforming is robust against passive eavesdropping in massive MIMO system [30]. But two oriented beamforming vectors are designed when the PSA is exposed to S, where one for suspicious messages that directed to D and another for interference signals that directed to E. According to the results of detection, there are four cases during data transmission phase. Specifically, **Case I:** S **is in favor of** H_0 **and** H_0 **is true**. **Case II:** S **is in favor of** H_0 **but** H_1 **is true**. **Case III:** S **is in favor of** H_1 **but** H_0 **is true**. **Case IV:** S **is in favor of** H_1 **and** H_1 **is true**. In this subsection, we will discuss these four cases, in detail.

Case I: During the channel training phase, S can estimate the CSI based on the received pilot symbols by the least-squares approach, i.e. $\hat{\mathbf{h}} = \mathbf{y}x^H$. For this case, the channel estimation result of $\hat{\mathbf{h}}_{SD}$ is

$$\hat{\mathbf{h}}_{SD} = \sqrt{P_D}\mathbf{h}_{SD} + \mathbf{n}x^H \tag{16}$$

And the MRT beamforming vector can be expressed as

$$\mathbf{w}_1 = \frac{\hat{\mathbf{h}}_{SD}}{\left\|\hat{\mathbf{h}}_{SD}\right\|} = \frac{\sqrt{P_D}\mathbf{h}_{SD}+\mathbf{n}x^H}{\left\|\sqrt{P_D}\mathbf{h}_{SD}+\mathbf{n}x^H\right\|} \tag{17}$$

In addition, to successfully eavesdrop the suspicious messages, the legitimate eavesdropper also sends interference signals through antenna J during the data transmission phase. By considering imperfect self-interference cancelation at E, the received suspicious signals at D and E can be respectively expressed as

$$y_D = \sqrt{P_S}\mathbf{h}_{SD}^H\mathbf{w}_1 s + \sqrt{P_J}h_{JD}\mathrm{z} + n \tag{18}$$

$$y_E = \sqrt{P_S}\mathbf{h}_{SE}^H\mathbf{w}_1 s + \sqrt{P_J}\phi h_{JD}\mathrm{z} + n \tag{19}$$

where P_S is the data transmission power, s denotes the suspicious data transmitted by S, which satisfied $E\left\{ss^H\right\} = 1$, $\phi \in (0, 1)$ denotes the coefficient of residual SI, P_J is jamming power, z denotes the jamming signal that satisfied $z \sim \mathcal{CN}(0, 1)$, and $n \sim \mathcal{CN}\left(0, \sigma_0^2\right)$ represents the received noise.

Case II: For this case, the channel estimation result of $\hat{\mathbf{h}}_{SD}$ can be computed as

$$\hat{\mathbf{h}}_{SD} = \sqrt{P_D}\mathbf{h}_{SD} + \sqrt{P_E}\mathbf{h}_{SE} + \mathbf{n}x^H \tag{20}$$

And the MRT beamforming vector is given as

$$\mathbf{w}_2 = \frac{\hat{\mathbf{h}}_{SD}}{\left\|\hat{\mathbf{h}}_{SD}\right\|} = \frac{\sqrt{P_D}\mathbf{h}_{SD}+\sqrt{P_E}\mathbf{h}_{SE}+\mathbf{n}x^H}{\left\|\sqrt{P_D}\mathbf{h}_{SD}+\sqrt{P_E}\mathbf{h}_{SE}+\mathbf{n}x^H\right\|} \tag{21}$$

The received suspicious signals at D and E can be respectively expressed as

$$y_D = \sqrt{P_S}\mathbf{h}_{SD}^H\mathbf{w}_2 s + \sqrt{P_J}h_{JD}\mathrm{z} + n \tag{22}$$

$$y_E = \sqrt{P_S}\mathbf{h}_{SE}^H\mathbf{w}_2 s + \sqrt{P_J}\phi h_{JD}\mathrm{z} + n \tag{23}$$

Case III: For this case, S is false alarm and the channel estimation result of $\hat{\mathbf{h}}_{SD}$ is

$$\hat{\mathbf{h}}_{SD} = \sqrt{P_D}\mathbf{h}_{SD} + \mathbf{n}x^H \tag{24}$$

Except $\hat{\mathbf{h}}_{SD}$, S deems that there is an extra wiretap channel, but there is not, in fact. We assume that S transmits interference signals with a random direction, which has nothing to do with $\mathbf{h}_{SE}$. So the MRT beamforming vectors are given as

$$\mathbf{w}_{31} = \frac{\hat{\mathbf{h}}_{SD}}{\left\|\hat{\mathbf{h}}_{SD}\right\|} = \frac{\sqrt{P_D}\mathbf{h}_{SD}+\mathbf{n}x^H}{\left\|\sqrt{P_D}\mathbf{h}_{SD}+\mathbf{n}x^H\right\|} \tag{25}$$

$$\mathbf{w}_{32} = \frac{\mathbf{n}x^H}{\left\|\mathbf{n}x^H\right\|} \tag{26}$$

For this case, the received suspicious signals at D and E can be respectively expressed as

$$y_D = \sqrt{P_{S1}}\mathbf{h}_{SD}^H\mathbf{w}_{31}s + \sqrt{P_{S2}}\mathbf{h}_{SD}^H\mathbf{w}_{32}a + \sqrt{P_J}h_{JD}\mathrm{z} + n \tag{27}$$

$$y_E = \sqrt{P_{S1}}\mathbf{h}_{SE}^H\mathbf{w}_{31}s + \sqrt{P_{S2}}\mathbf{h}_{SE}^H\mathbf{w}_{32}a + \sqrt{P_J}\phi h_{JE}\mathrm{z} + n \tag{28}$$

where $P_{S1} + P_{S2} = P_S$ and a denotes the interference signal transmitted by S and satisfied that $E\left\{aa^H\right\} = 1$.

Case VI: For this case, the PSA is exposed to S, we assume that S can distinguish the pilot symbols through an extra process [31], [32]. Then, the channel estimation results of $\mathbf{h}_{SD}$ and $\mathbf{h}_{SE}$ under this case are

$$\hat{\mathbf{h}}_{SD} = \sqrt{P_D}\mathbf{h}_{SD} + \mathbf{n}x^H \tag{29}$$

$$\hat{\mathbf{h}}_{SE} = \sqrt{P_E}\mathbf{h}_{SE} + \mathbf{n}x^H \tag{30}$$

The beamforming vectors are given as

$$\mathbf{w}_{41} = \frac{\hat{\mathbf{h}}_{SD}}{\left\|\hat{\mathbf{h}}_{SD}\right\|} = \frac{\sqrt{P_D}\mathbf{h}_{SD} + \mathbf{n}x^H}{\left\|\sqrt{P_D}\mathbf{h}_{SD} + \mathbf{n}x^H\right\|} \tag{31}$$

$$\mathbf{w}_{42} = \frac{\hat{\mathbf{h}}_{SE}}{\left\|\hat{\mathbf{h}}_{SE}\right\|} = \frac{\sqrt{P_E}\mathbf{h}_{SE} + \mathbf{n}x^H}{\left\|\sqrt{P_E}\mathbf{h}_{SE} + \mathbf{n}x^H\right\|} \tag{32}$$

For this case, the received suspicious messages at D and E can be respectively expressed as

$$y_D = \sqrt{P_{S1}}\mathbf{h}_{SD}^H\mathbf{w}_{41}s + \sqrt{P_{S2}}\mathbf{h}_{SD}^H\mathbf{w}_{42}a + \sqrt{P_J}h_{JD}\mathrm{z} + n \tag{33}$$

$$y_E = \sqrt{P_{S1}}\mathbf{h}_{SE}^H\mathbf{w}_{41}s + \sqrt{P_{S2}}\mathbf{h}_{SE}^H\mathbf{w}_{42}a + \sqrt{P_J}\phi h_{JE}\mathrm{z} + n \tag{34}$$

Remark 2: It is worth to note that Case I is similar as the jamming scheme proposed in [15], and the main difference is that the suspicious source equips with multi-antennas in our case. Case II and Case III are two possible cases of proposed covert PSA scheme, where Case II indicates missed detection and Case II means false alarm. Case VI is a worst case that the PSA is exposed to suspicious source.

IV. EAVESDROPPING RATE ANALYSIS AND OPTIMIZATION

A. AVERAGE EFFECTIVE EAVESDROPPING RATE ANALYSIS

In this section, we derive the closed-form expression of effective eavesdropping rate according to the four cases and optimize the power allocation at eavesdropper for the proposed covert PSA scheme.

Case I: According to the received suspicious signals (18) and (19) during the data transmission phase, the average data rates at D and E can be respectively computed as

$$R_D = E_{\mathbf{h}_{SD},h_{JD}}\left\{\log_2\left(1 + \frac{P_S\left\|\mathbf{h}_{SD}^H\mathbf{w}_1\right\|^2}{P_J|h_{JD}|^2 + \sigma_0^2}\right)\right\} \approx \log_2\left(1 + \frac{P_S\beta_{JD}\left(P_DN_S + \beta_{SD}\sigma_0^2\right)}{\beta_{SD}\left(P_D + \beta_{SD}\sigma_0^2\right)\left(P_J + \beta_{JD}\sigma_0^2\right)}\right) \tag{35}$$

$$R_E = E_{\mathbf{h}_{SD},\mathbf{h}_{SE},h_{JE}}\left\{\log_2\left(1 + \frac{P_S\left\|\mathbf{h}_{SE}^H\mathbf{w}_1\right\|^2}{P_J\phi^2|h_{JE}|^2 + \sigma_0^2}\right)\right\} \approx \log_2\left(1 + \frac{P_S\beta_{JE}\left(P_D + \beta_{SD}\sigma_0^2\right)}{\beta_{SE}\left(P_D + \beta_{SD}\sigma_0^2\right)\left(P_J\phi^2 + \beta_{JE}\sigma_0^2\right)}\right) \tag{36}$$

Proof: See Appendix A.

Case II: Similarly, according to (22) and (23) during the data transmission phase, the average data rates of this case at D and E can be respectively computed as

$$R_D = E_{\mathbf{h}_{SD},\mathbf{h}_{SE},h_{JD}}\left\{\log_2\left(1 + \frac{P_S\left\|\mathbf{h}_{SD}^H\mathbf{w}_2\right\|^2}{P_J|h_{JD}|^2 + \sigma_0^2}\right)\right\} \approx \log_2\left(1 + \frac{P_S\beta_{JD}}{P_J + \beta_{JD}\sigma_0^2}\frac{\frac{N_SP_D}{\beta_{SD}^2} + \frac{P_E}{\beta_{SD}\beta_{SE}} + \frac{\sigma_0^2}{\beta_{SD}}}{\frac{P_D}{\beta_{SD}} + \frac{P_E}{\beta_{SE}} + \sigma_0^2}\right) \tag{37}$$

$$R_E = E_{\mathbf{h}_{SD},\mathbf{h}_{SE},h_{JE}}\left\{\log_2\left(1 + \frac{P_S\left\|\mathbf{h}_{SE}^H\mathbf{w}_2\right\|^2}{P_J\phi^2|h_{JE}|^2 + \sigma_0^2}\right)\right\} \approx \log_2\left(1 + \frac{P_S\beta_{JE}}{P_J\phi^2 + \beta_{JE}\sigma_0^2}\frac{\frac{N_SP_E}{\beta_{SE}^2} + \frac{P_D}{\beta_{SD}\beta_{SE}} + \frac{\sigma_0^2}{\beta_{SE}}}{\frac{P_D}{\beta_{SD}} + \frac{P_E}{\beta_{SE}} + \sigma_0^2}\right) \tag{38}$$

Case III: Furthermore, according to (27) and (28) during the data transmission phase, the average data rates of this case at D and E can be respectively computed as (39) and (40), which is shown at the bottom of the next page.

Case VI: In addition, according to (33) and (34) during the data transmission phase, the average data rates of this case at D and E can be respectively computed as (41) and (42), which is also shown at the bottom of the next page.

Different from the existing works [14]–[18] where the legitimate eavesdropper optimizes the power allocation based on instantaneous CSI, we define the effective eavesdropping rate as a metric which the legitimate eavesdropper can select PSA and jamming power according to the statistical CSI to maximize the legitimate monitoring performance. Note that the effective eavesdropping rate can also approximate the exact legitimate eavesdropping performance well. Specifically, for the case that the PSA is covered, the eavesdropper can successfully eavesdrop the suspicious messages whit high probability (approximately equal to 1), since some message beams are directed to legitimate eavesdropper and the active jamming can further worse the received SINR at suspicious destination. And for the case that the PSA is exposed to suspicious source, the eavesdropper can hardly eavesdrop the suspicious messages (approximately equal to 0) as the artificial noise beams are pointed to legitimate eavesdropper but the message beams are directed to suspicious destination. Hence, for the considered proactive eavesdropping system, the effective eavesdropping rate can be defined as

$$R_M = \begin{cases} R_D & R_E \geq R_D \\ 0 & R_E < R_D \end{cases} \tag{43}$$

Remark 3: Based on the closed-form expression of average data rate in Case I, the challenge to successfully eavesdrop a suspicious communication system that equipped with multi-antenna is revealed. With the number of antennas N_S increases, R_D increases but R_E remains unchanged, hence, even J jams with the maximum transmission power, R_E may be still less than R_D. For case II, the PSA is missed detection and some suspicious messages are directed to E, while some power of S is allocated to transmit interference with a random orientation for case III, which has more affect on D with respect to the received signal to interference plus noise ratio (SINR) since the performance of cognitive jamming becomes better with P_S decreasing. Case VI lists a worst case that the PSA is exposed to S. For this case, it's more challenge to eavesdrop suspicious messages as the messages beam is directed to D while the interference beam is directed to E. So, it's necessary to launch pilot attack while keeping it covert to S.

Remark 4: It's obvious that the effective eavesdropping rate highly depends on the power allocation between pilot and interference at eavesdropper. The more power to transmit pilot can accelerate suspicious messages leakage but is easy to be exposed. Similarly, the more power to send jamming can ensure a positive eavesdropping rate with high probability but the effective eavesdropping rate may also be degraded. So, it's necessary to optimize the power allocation for eavesdropper.

B. OPTIMAL POWER ALLOCATION AT EAVESDROPPER

In this subsection, we optimize the transmission power of legitimate eavesdropper to maximize the effective eavesdropping rate in the considered legitimate eavesdropping system under the covert PSA and transmission power constraints. Specifically, the optimization problem can be formulated as

$$\begin{aligned} &\max_{P_E,P_J} \quad R_M \\ &s.t.\ C1: \xi^* \geq 1-\varepsilon \\ &\qquad C2: 0 \leq P_E + P_J \leq P_{\max} \end{aligned} \tag{44}$$

where $C1$ is the covert PSA requirement that ε is predetermined covert constraint and $C2$ is the transmission power budgets for legitimate eavesdropper.

For $R_E < R_D$, the effective eavesdropping rate is zero, so the optimal transmission power is $P_E = P_J = 0$. For $R_E \geq R_D$, by substituting (18) in (19), the optimal problem can be transformed as

$$\begin{aligned} &\max_{P_E,P_J} \quad R_D \\ &s.t.\ C1: \xi^* \geq 1-\varepsilon \\ &\qquad C2: 0 \leq P_E + P_J \leq P_{\max} \\ &\qquad C3: R_E \geq R_D \end{aligned} \tag{45}$$

To solve such a non-convex problem, we firstly treat P_E as a constant while optimizing the jamming power P_J.

$$\begin{aligned} Case\ III:\ R_D &= E_{\mathbf{h}_{SD},\mathbf{h}_{SE},h_{JD}}\left\{\log_2\left(1+\frac{P_{S1}\left\|\mathbf{h}_{SD}^H\mathbf{w}_{31}\right\|^2}{P_{S2}\left\|\mathbf{h}_{SD}^H\mathbf{w}_{32}\right\|^2+P_J|h_{JD}|^2+\sigma_0^2}\right)\right\} \\ &\approx \log_2\left(1+\frac{P_{S1}P_DN_S+P_{S1}\beta_{SD}\sigma_0^2}{P_D\beta_{SD}+\sigma_0^2\beta_{SD}^2}\frac{\beta_{JD}\beta_{SD}}{P_{S2}\beta_{JD}+P_J\beta_{SD}+\sigma_0^2\beta_{SD}\beta_{JD}}\right) \end{aligned} \tag{39}$$

$$\begin{aligned} R_E &= E_{\mathbf{h}_{SD},\mathbf{h}_{SE},h_{JE}}\left\{\log_2\left(1+\frac{P_{S1}\left\|\mathbf{h}_{SE}^H\mathbf{w}_{31}\right\|^2}{P_{S2}\left\|\mathbf{h}_{SE}^H\mathbf{w}_{32}\right\|^2+\phi^2P_J|h_{JE}|^2+\sigma_0^2}\right)\right\} \\ &\approx \log_2\left(1+\frac{P_{S1}P_D+P_{S1}\beta_{SD}\sigma_0^2}{P_D\beta_{SE}+\sigma_0^2\beta_{SE}\beta_{SD}}\frac{\beta_{JE}\beta_{SE}}{P_{S2}\beta_{JE}+P_J\phi^2\beta_{SE}+\beta_{SE}\beta_{JE}\sigma_0^2}\right) \end{aligned} \tag{40}$$

$$\begin{aligned} &Case\ VI:\ R_D \\ &= E_{\mathbf{h}_{SD},\mathbf{h}_{SE},h_{JD}}\left\{\log_2\left(1+\frac{P_{S1}\left\|\mathbf{h}_{SD}^H\mathbf{w}_{41}\right\|^2}{P_{S2}\left\|\mathbf{h}_{SD}^H\mathbf{w}_{42}\right\|^2+P_J|h_{JD}|^2+\sigma_0^2}\right)\right\} \\ &\approx \log_2\left(1+\frac{P_{S1}P_DN_S+P_{S1}\beta_{SD}\sigma_0^2}{\beta_{SD}P_D+\beta_{SD}^2\sigma_0^2}\frac{\beta_{JD}\left(\beta_{SD}+\beta_{SE}\beta_{SD}\sigma_0^2\right)}{P_{S2}\beta_{JD}\left(P_E+\beta_{SE}\sigma_0^2\right)+P_J\left(\beta_{SD}+\beta_{SE}\beta_{SD}\sigma_0^2\right)+\left(\beta_{SD}+\beta_{SE}\beta_{SD}\sigma_0^2\right)\beta_{JD}\sigma_0^2}\right) \end{aligned} \tag{41}$$

$$\begin{aligned} R_E &= E_{\mathbf{h}_{SD},\mathbf{h}_{SE},h_{JE}}\left\{\log_2\left(1+\frac{P_{S1}\left\|\mathbf{h}_{SE}^H\mathbf{w}_{41}\right\|^2}{P_{S2}\left\|\mathbf{h}_{SE}^H\mathbf{w}_{42}\right\|^2+\phi^2P_J|h_{JE}|^2+\sigma_0^2}\right)\right\} \\ &\approx \log_2\left(1+\frac{P_{S1}P_D+P_{S1}\beta_{SD}\sigma_0^2}{P_D\beta_{SE}+\beta_{SE}\beta_{SD}\sigma_0^2}\frac{\beta_{JE}\left(P_E\beta_{SE}+\beta_{SE}^2\sigma_0^2\right)}{P_{S2}\beta_{JE}\left(P_EN_S+\beta_{SE}\sigma_0^2\right)+P_J\phi^2\left(P_E\beta_{SE}+\beta_{SE}^2\sigma_0^2\right)+\left(P_E\beta_{SE}+\beta_{SE}^2\sigma_0^2\right)\beta_{JE}\sigma_0^2}\right) \end{aligned} \tag{42}$$

The problem (45) can be rewritten as

$$\begin{aligned} &\max_{P_J} \quad R_D \\ &s.t.\ C2': 0 \le P_J \le P_{\max} - P_E \\ &\quad C3: R_E \ge R_D \end{aligned} \tag{46}$$

For a given P_E, the constraint $C3$ can be transformed to $C3'$, which is

$$C3': P_J \ge \max\left\{\frac{\beta_{SD}\beta_{SE}\sigma_0^2\,(l_1 - l_2)}{\beta_{JE} l_2 - \beta_{JD}\phi^2 l_1}, 0\right\} \tag{47}$$

proof: See Appendix B.

As the interference transmitted by J simultaneously jams D and E, R_D and R_E monotonically decrease with P_J increaseing, which indicates that a optimal jamming power of J is the minimum one that satisfied $R_E \ge R_D$. Hence, a optimal jamming power during the data transmission phase can be computed as

$$P_J{}^* = \max\left\{\frac{\beta_{SD}\beta_{SE}\sigma_0^2\,(l_1 - l_2)}{\beta_{JE} l_2 - \beta_{JD}\phi^2 l_1}, 0\right\} \tag{48}$$

After obtaining the optimal jamming power $P_J{}^*$, the problem to optimize P_E can be formed as

$$\begin{aligned} &\max_{P_E} \quad R_D \\ &s.t.\ C1: \xi^* \ge 1 - \varepsilon \\ &\quad C2'': 0 \le P_E + P_J{}^* \le P_{\max} \end{aligned} \tag{49}$$

With respect to P_E, it can be solved via one-dimensional linear search. **Algorithm** 1 is summarized to solve problem (49). As to the complexity of our proposed algorithm, it highly depends on the cycles of one-dimensional search as each cycle only involves numerical operations. If we define the complexity of each numerical operations in the cycle as T_c, the total computational complexity is $O(\frac{T_C P_{\max}}{\|\Delta\|})$, as the maximum number of iterations is $\frac{P_{\max}}{\|\Delta\|}$.

Algorithm 1 Optimal Power Allocation Algorithm

1: Set the initial power $P_E^{(0)}$, maximal transmission power $P_{\max}$, step length Δ and error margin ε.
2: **for** $l = 1, 2, \cdots$ **do**
3: **if** $\xi \ge 1 - \varepsilon$.
4: compute P_J^* according to (48).
5: **if** $P_J^* \le P_{\max} - P_E{}^{(k)}$.
6: compute $R_D{}^{(k)}$ according to (37).
7: $P_E{}^{(k+1)} = P_E{}^{(k)} + \Delta$.
8: **end for until** $P_E{}^{(k)} \le 0$ or $P_E{}^{(k)} \ge P_{\max}$.
9: Obtain the maximal R_D and optimal P_E^*, P_J^*.

V. SIMULATION RESULTS

In this section, we first present numerical results to illustrate the detection performance at S for the proposed covert PSA scheme and demonstrate the optimal power allocation algorithm of the legitimate eavesdropper to maximize the effective eavesdropping rate is feasible. In our simulations, without loss of generality, we set the large scale fading coefficients as $\beta_{SD} = \beta_{SE} = \beta_{JD} = 1$ and the power of noise is $\sigma_0^2 = 0.1\ W$. The power of pilot at D and the power of suspicious messages at S are equal and fixed, i.e., $P_D = P_S = 2\ W$. Furthermore, we assume that half of power at S is allocated to send interference in case III and case VI, while all power is used to send data in case I and II. Also, the power at legitimate eavesdropper is equally allocated for the antenna E and J to transmit pilot symbol and interference during training and data transmission phase, respectively.

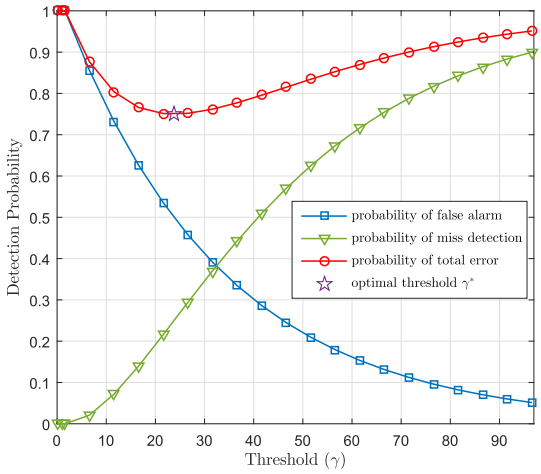

FIGURE 3. Detection probability versus detection threshold.

Fig. 3 depicts the probability of false alarm p_{fa}, the probability of missed detection p_{md}, the probability of total error ξ versus the detection threshold γ where $N_S = 16$ and $P_E = 0.5\ W$. As our theoretical analysis, the probability of false alarm is 1, while the probability of missed detection keeps 0 when the threshold of detection is less than $N_S\sigma_0^2 = 1.6$. And the probability of false alarm decreases while the probability of missed detection increases with the threshold increasing when the threshold is lager than $N_S\sigma_0^2$. As a result, the probability of total error also is 1 when the threshold of detection is less than $N_S\sigma_0^2$ but first decreases and then increases when the threshold is lager than $N_S\sigma_0^2$, which indicates that there is a optimal detection threshold for detector. However, $\xi > 0$ is true even under the optimal detection threshold, which means that uncertainty of channels can be utilized to achieve covert PSA for legitimate eavesdropper.

Fig. 4 plots the probability of total error ξ versus the detection threshold γ with different P_E and N_S. In this figure, we first observe that the optimal threshold of detection increases with P_E and N_S increase, respectively. And we can see the minimal ξ decreases when P_E increases, which means that the PSA is easier to be detected by S when the P_E increases. So, the available power of PSA is finite as the covert constraint normally requires $\xi > 1 - \varepsilon$, for any $\varepsilon > 0$. Although the proposed covert PSA scheme can be adopted to assist legitimate monitoring during the channel

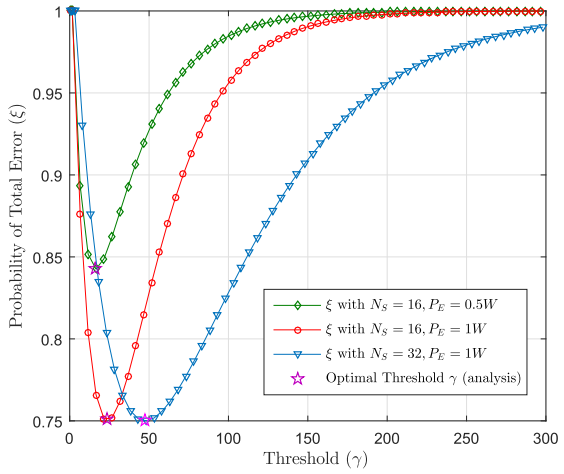

FIGURE 4. Total error detection probability versus detection threshold.

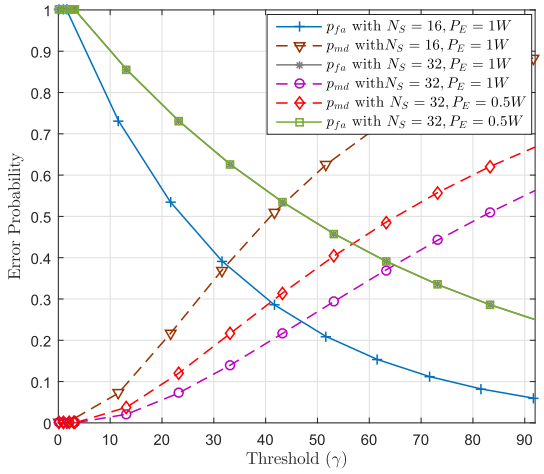

FIGURE 5. Detection probability versus detection threshold.

training phase, the active jamming is also necessary since P_E is less than P_D in general and most power of suspicious message beam is directed to D. In addition, we observe that the minimal ξ remains unchanged when N_S increases, which indicates that the probability of total error has nothing to do with N_S under a optimal detection threshold at S.

Fig. 5 plots the probability of missed detection p_{md} and false alarm p_{fa} versus the detection threshold γ with P_E and N_S. In this figure, we can see that p_{md} increases with P_E and N_S decreasing, respectively. Hence, for the case that S equipped with massive antennas, the PSA may be detected with a high probability even the power of PSA is small. However, p_{fa} decreases with N_S increasing but keeps unchanged when P_E increases, so the total error detection may still satisfy the covert constraint when the number of antennas increase.

Fig. 6 shows the discriminatory of achievable rate R_E and R_D versus the number of antenna N_S with different cases during data transmission phase, where we set $P_E = P_J = 1W$ and the SI coefficient is $\phi = 0.1$. From the figure, we can see that the discriminatory of achievable rate $R_E - R_D$ slowly

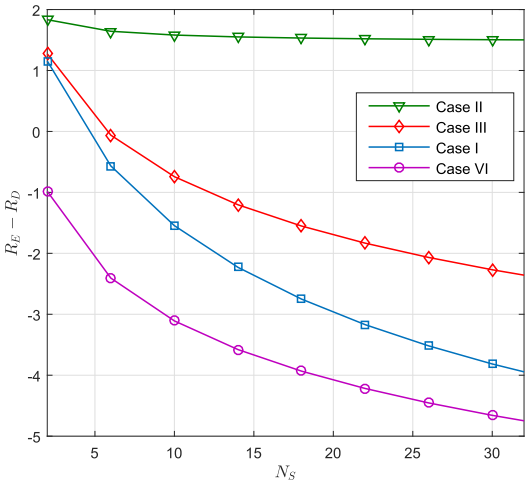

FIGURE 6. Discriminatory of achievable rate R_E and R_D versus the number of antenna N_S.

decreases and even remains constant in case II when the number of antenna at S is large enough. Furthermore, $R_E - R_D$ is more than zero for this case, which indicates that legitimate monitoring is effective with the proposed legitimate monitoring scheme even in a multiple antenna system since some suspicious information beams are directly leaked to legitimate eavesdropper and most SI can be effectively eliminated. For comparing, we can observe that $R_E - R_D$ persistently decreases with N_S increasing for case I, case III and case VI, which leads legitimate monitoring invalid ($R_E - R_D < 0$) when N_S is larger than a threshold. For case III, although some power is allocated to send interference with a random orientation and SINR is more seriously worsen, there is a tiny improvement by comparing to case I. For case VI, since the the PSA is exposed to S and the interference is directly pointed to E via beamforming, it has a worst legitimate monitoring performance. It's also revealed that the PSA can effectively assist to monitor a suspicious multiple antenna system but the PSA needs to be tactfully hidden without being detected for limited transmission power at eavesdropper.

Fig. 7 depicts the discriminatory of achievable rate R_E and R_D versus the SI coefficient ϕ with different N_S and cases, where $P_E = P_J = 1W$. We can see that $R_E - R_D$ persistently decreases when ϕ increases since R_E decreases with ϕ increasing. When ϕ is larger than 0.5, the legitimate monitoring becomes invalid even with the help of covert PSA. And for Case I, even though $\phi = 0$, which corresponds to a perfected cancellation of SI, the legitimate monitoring may still be invalid as $R_E - R_D < 0$. It also shows that the number of antenna N_S has little affect on the proposed covert PSA scheme, while active jamming scheme proposed in [15] is severely affected by N_S.

Fig. 8 shows the effective eavesdropping rate versus the number of antennas N_S. We can see that the effective eavesdropping rate slowly decreases and then rapidly decreases to 0 when the number of antennas at S increases for the proactive eavesdropping without PSA, however, it slowly increases and

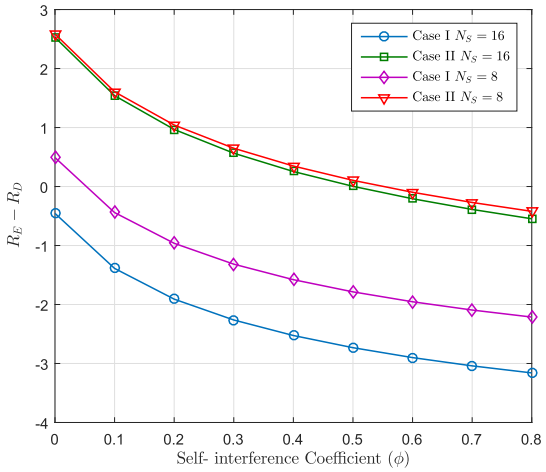

FIGURE 7. Discriminatory of achievable rate R_E and R_D versus the self-interference coefficient ϕ.

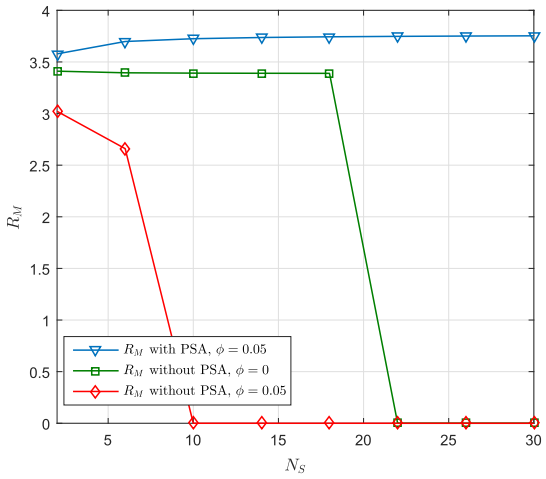

FIGURE 8. Effective eavesdropping rate versus the number of antennas N_S.

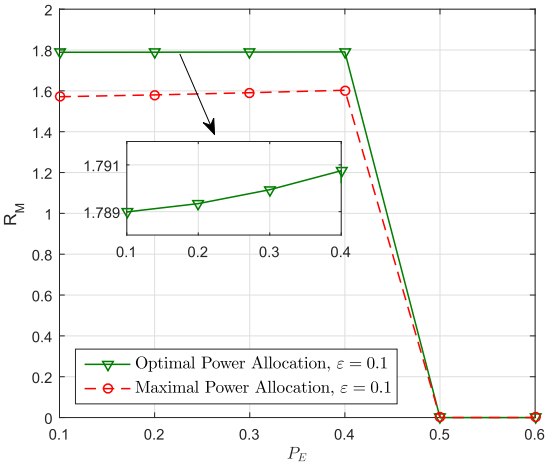

FIGURE 9. Effective eavesdropping rate versus the power of PSA.

then maintains for the proposed covert PSA scheme. This is because that $R_E < R_D$ even all the power are used to transmit jamming for the proactive eavesdropping without PSA when the number of antennas at S is lager than a threshold. But for the proposed covert PSA scheme, it's robust against multi-antenna suspicious communication since some suspicious messages beams are directed to the legitimate eavesdropper.

Fig. 9 plots the effective eavesdropping rate versus the power allocation at eavesdropper. In this figure, we first observe that the effective eavesdropping rate slowly increases and then rapidly decreases to zero with the power of PSA increasing, this is due to the fact that more suspicious messages leak to eavesdropper with the power of PSA increasing, however, the PSA may be exposed to S (i.e. $\xi^* < 1 - \varepsilon$) when the power of PSA is lager than a threshold, which leads to unsuccessfully eavesdropping. Furthermore, we observe that the proposed power allocation scheme can consume less power while achieving larger effective eavesdropping rate by comparing to the maximal power allocation scheme. This is because that, on the one hand, more power to transmit jamming will decrease the pilot attack power as the total power is constraints, on the other hand, the active jamming also decreases the performance at E since the imperfect interference elimination.

VI. CONCLUSION

In this paper, we first analyze the total error detection performance of suspicious source and derive the close-form expression of optimal detection threshold and effective eavesdropping rate. Then, we optimize the power allocation between PAS and interference to maximize the effective eavesdropping rate with the covert PSA and transmission power constraints. Finally, the numerical results are provided to validate the analysis. In addition, we show that the uncertainty of channels during the training phase can be utilized to achieve covert PSA for legitimate eavesdropper. Furthermore, the proposed covert PSA scheme can effectively combat with multi-antenna suspicious communication system.

APPENDIX A

According to the following approximation [43]

$$E\left\{\log_2\left(1+\frac{X}{Y}\right)\right\} \approx \log_2\left(1+\frac{E\{X\}}{E\{Y\}}\right) \tag{50}$$

we can formulate the average rate at D as

$$\begin{aligned} R_D &\approx \log_2\left(1+\frac{E\left\{P_S\left\|\mathbf{h}_{SD}^H\mathbf{w}_1\right\|^2\right\}}{E\left\{P_J|h_{JD}|^2\right\}+\sigma_N^2}\right) \\ &\approx \log_2\left(1+\frac{P_S\beta_{JD}}{P_J+\beta_{JD}\sigma_N^2}\frac{E\left\{\left\|\mathbf{h}_{SD}^H\hat{\mathbf{h}}_{SD}\right\|^2\right\}}{E\left\{\left\|\hat{\mathbf{h}}_{SD}\right\|^2\right\}}\right) \end{aligned} \tag{51}$$

Substituting $\hat{\mathbf{h}}_{SD} = \sqrt{P_D}\mathbf{h}_{SD} + \mathbf{n}x^H$ into (51), we can respectively compute the $E\left\{\left\|\mathbf{h}_{SD}^H\hat{\mathbf{h}}_{SD}\right\|^2\right\}$ and $E\left\{\left\|\hat{\mathbf{h}}_{SD}\right\|^2\right\}$

$$\frac{\beta_{JE}}{P_J\phi^2+\beta_{JE}\sigma_0^2}\left(\frac{N_S P_E}{\beta_{SE}^2}+\frac{P_D}{\beta_{SD}\beta_{SE}}+\frac{\sigma_0^2}{\beta_{SE}}\right) > \frac{\beta_{JD}}{P_J+\beta_{JD}\sigma_0^2}\left(\frac{N_S P_D}{\beta_{SD}^2}+\frac{P_E}{\beta_{SD}\beta_{SE}}+\frac{\sigma_0^2}{\beta_{SD}}\right) \tag{54}$$

as

$$\begin{aligned} E_{\mathbf{h}_{SD}}\left\{\left\|\mathbf{h}_{SD}^H\hat{\mathbf{h}}_{SD}\right\|^2\right\} &= E_{\mathbf{h}_{SD}}\left(P_D\mathbf{h}_{SD}^H\mathbf{h}_{SD}\mathbf{h}_{SD}^H\mathbf{h}_{SD}\right) \\ &\quad + E_{\mathbf{h}_{SD}}\left(\mathbf{h}_{SD}^H\mathbf{n}_D x^H x\mathbf{n}_D^H\mathbf{h}_{SD}\right) \\ &= \frac{P_D(N_S)^2}{\beta_{SD}^2}+\frac{\sigma_N^2 N_S}{\beta_{SD}} \end{aligned} \tag{52}$$

$$\begin{aligned} E_{\mathbf{h}_{SD}}\left\{\left\|\hat{\mathbf{h}}_{SD}\right\|^2\right\} &= E_{\mathbf{h}_{SD}}\left(P_D\mathbf{h}_{SD}^H\mathbf{h}_{SD}\right)+E\left(\mathbf{n}_D x^H x\mathbf{n}_D^H\right) \\ &= \frac{P_D N_S}{\beta_{SD}}+\sigma_N^2 N_S \end{aligned} \tag{53}$$

By substituting (52) and (53) into (51), we can obtain the average data rates at D as (35). Similarly, we can obtain the proof of (36).

Since the random variable X and Y in approximation expression (50) are not required to be independent [43], we can demonstrate the approximation expression of achievable rate in case II, III and VI, respectively.

APPENDIX B

By substituting (37) and (38) into $C3$, we can simplify the $C3$ as (54), which is shown at the top of this page. By defining

$$l_1 = P_D N_S\beta_{SE}^2+P_E\beta_{SD}\beta_{SE}+\sigma_0^2\beta_{SE}^2\beta_{SD} \tag{55}$$

$$l_2 = P_E N_S\beta_{SD}^2+P_D\beta_{SD}\beta_{SE}+\sigma_0^2\beta_{SD}^2\beta_{SE} \tag{56}$$

(54) can be further computed as

$$P_J\left(\beta_{JE}l_2-\beta_{JD}\phi^2 l_1\right) > \beta_{JE}\beta_{JD}\sigma_0^2\left(l_1-l_2\right) \tag{57}$$

If l_1 and l_2 satisfy that $l_1 < l_2$, which means the pilot power received from E is larger than that from D, the optimal interference power is $P_J = 0$ as most energy of suspicious message beam is directed to E and there is no need to send jamming. Otherwise, the interference power should be satisfied that $P_J \geq \frac{\beta_{SD}\beta_{SE}\sigma_0^2(l_1-l_2)}{\beta_{JE}l_2-\beta_{JD}\phi^2 l_1}$ when $\beta_{JE}l_2-\beta_{JD}\phi^2 l_1 > 0$. In addition, if l_1 and l_2 satisfy that $l_1 > l_2$ and $\beta_{JE}l_2-\beta_{JD}\phi^2 l_1 < 0$, which indicates that the pilot power received from E is less than that from D and the there is a serious SI, the constraint $C3$ is never satisfied for any P_J. Hence, constraint $C3$ can be transformed as $C3'$ that given as (47).

REFERENCES

[1] S. Mansoor, A. Molisch, P. Smith, T. Haustein, and G. Wunder, "5G: A tutorial overview of standards, trials, challenges, deployment and practice," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1201–1221, Jun. 2017.

[2] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1617–1655, 3rd Quart., 2016.

[3] Y. Gu, Z. Wu, Z. Yin, and X. Zhang, "The secrecy capacity optimization artificial noise: A new type of artificial noise for secure communication in MIMO system," *IEEE Access*, vol. 7, pp. 58353–58360, 2019.

[4] N. Zhao, F. R. Yu, M. Li, Q. Yan, and V. C. M. Leung, "Physical layer security issues in interference-alignment-based wireless networks," *IEEE Commun. Mag.*, vol. 54, no. 8, pp. 162–168, Aug. 2016.

[5] N. Zhao, Y. Cao, F. R. Yu, Y. Chen, M. Jin, and V. C. M. Leung, "Artificial noise assisted secure interference networks with wireless power transfer," *IEEE Trans. Wireless Commun.*, vol. 67, no. 2, pp. 1087–1098, Feb. 2018.

[6] Y. Feng, S. Yan, Z. Yang, N. Yang, and J. Yuan, "Beamforming design and power allocation for secure transmission with NOMA," *IEEE Trans. Wireless Commun.*, vol. 18, no. 5, pp. 2639–2651, May 2019.

[7] J. Xiong, D. Ma, K.-K. Wong, and J. Wei, "Robust masked beamforming for MISO cognitive radio networks with unknown eavesdroppers," *IEEE Trans. Veh. Technol.*,vol. 65, no. 2, pp. 744–755, Feb. 2016.

[8] L. Hu, H. Wen, B. Wu, F. Pan, R. F. Liao, H. Song, J. Tang, and X. Wang, "Cooperative jamming for physical layer security enhancement in Internet of Things," *IEEE Internet Things J.*, vol. 5, no. 1, pp. 219–228, Feb. 2018.

[9] Y. Liu, H.-H. Chen, and L. Wang, "Physical layer security for next generation wireless networks: Theories, technologies, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 347–376, 1st Quart., 2017.

[10] Z. Xiang, W. Yang, G. Pan, Y. Cai, and Y. Song, "Physical layer security in cognitive radio inspired NOMA network," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 3, pp. 700–714, Jun. 2019.

[11] F. Cheng, G. Gui, N. Zhao, Y. Chen, J. Tang, and H. Sari, "UAV-relaying-assisted secure transmission with caching," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3140–3153, May 2019.

[12] T. Zhang, Y. Cai, Y. Huang, T. Q. Duong, and W. Yang, "Secure full-duplex spectrum-sharing wiretap networks with different antenna reception schemes," *IEEE Trans. Commun.*,vol. 65, no. 1, pp. 335–346, Jan. 2019.

[13] X. Sun, W. Yang, Y. Cai, L. Tao, Y. Liu, and Y. Huang, "Secure transmissions in wireless information and power transfer millimeter-wave ultra-dense networks," *IEEE Trans. Commun.*, vol. 14, no. 7, pp. 1817–1829, Jul. 2019.

[14] J. Xu, L. Duan, and R. Zhang, "Surveillance and intervention of infrastructure-free mobile communications: A new wireless security paradigm," *IEEE Wireless Commun.*, vol. 24, no. 4, pp. 152–159, Aug. 2017.

[15] X. Jie, L. Duan, and R. Zhang, "Proactive eavesdropping via cognitive jamming in fading channels," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2790–2806, May 2017.

[16] S. Huang, Q. Zhang, Q. Li, and J. Qin, "Robust proactive monitoring via jamming with deterministically bounded channel errors," *IEEE Signal Process. Lett.*, vol. 25, no. 5, pp. 690–694, May 2018.

[17] W. Huang, W. Chen, B. Bai, and Z. Han, "Wiretap channel with full-duplex proactive eavesdropper: A game theoretic approach," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7658–7663, Aug. 2018.

[18] Y. Cai, C. Zhao, Q. Shi, G. Y. Li, and B. Champagne, "Joint beamforming and jamming design for mmWave information surveillance systems," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 7, pp. 1410–1425, Jul. 2018.

[19] D. Xu, H. Zhu, and Q. Li, "Jammer-assisted legitimate eavesdropping in wireless powered suspicious communication networks," *IEEE Access*, vol. 7, pp. 20363–20380, 2019.

[20] H. Lu, H. Zhang, H. Dai, W. Wu, and B. Wang, "Proactive eavesdropping in UAV-aided suspicious communication systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1993–1997, Feb. 2019.

[21] Y. Zeng and R. Zhang, "Wireless information surveillance via proactive eavesdropping with spoofing relay," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 8, pp. 1449–1461, Dec. 2016.

[22] J. Moon, H. Lee, C. Song, S. Kang, and I. Lee, "Relay-assisted proactive eavesdropping with cooperative jamming and spoofing," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6958–6971, Oct. 2018.

[23] J. Moon, H. Lee, C. Song, S. Lee, and I. Lee, "Proactive eavesdropping with full-duplex relay and cooperative jamming," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6707–6719, Oct. 2018

[24] B. Li, Y. Yao, H. Zhang, and Y. Lv, "Energy efficiency of proactive cooperative eavesdropping over multiple suspicious communication links," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 420–430, Jan. 2019.

[25] H. Cai, Q. Zhang, Q. Li, and J. Qin, "Proactive monitoring via jamming for rate maximization over MIMO Rayleigh fading channels," *IEEE Commun. Lett.*, vol. 21, no. 9, pp. 2021–2024, Sep. 2017.

[26] 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, Jun. 2015.

[27] X. Zhou, B. Maham, and A. Hjorungnes, "Pilot contamination for active eavesdropping," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 903–907, Mar. 2012.

[28] K.-W. Huang, H.-M. Wang, Y. Wu, and R. Schober, "Pilot spoofing attack by multiple eavesdroppers," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6433–6447, Oct. 2018.

[29] L. Wan, G. Zhang, M. Cui, and F. Lin, "Proactive eavesdropping via pilot contamination and jamming," *Wireless Pers. Commun.*, vol. 99, no. 3, pp. 1405–1421, Apr. 2018.

[30] Q. Xiong, Y.-C. Liang, K. H. Li, and Y. Gong, "An energy-ratio-based approach for detecting pilot spoofing attack in multiple-antenna systems," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 5, pp. 932–940, May 2015.

[31] Q. I. Xiong, Y.-C. Liang, K. H. Li, Y. Gong, and S. Han, "Secure transmission against pilot spoofing attack: A two-way training-based scheme," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 5, pp. 1017–1026, May 2017.

[32] X. Tian, M. Li, and Q. Liu, "Random-training-assisted pilot spoofing detection and security enhancement," *IEEE Access*, vol. 5, pp. 27384–27399, 2017.

[33] T. V. Sobers, B. A. Bash, S. Guha, D. Towsley, and D. Goeckel, "Covert communication in the presence of an uninformed jammer," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6193–6206, Sep. 2017.

[34] B. He, S. Yan, X. Zhou, and H. Jafarkhani, "Covert wireless communication with a Poisson field of interferers," *IEEE Trans. Wireless Commun.*, vol. 17, no. 9, pp. 6005–6017, Sep. 2018.

[35] T.-X. Zheng, H.-M. Wang, D. W. K. Ng, and J. Yuan, "Multi-antenna covert communications in random wireless networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1974–1987, Mar. 2019.

[36] J. Hu, S. Yan, X. Zhou, F. Shu, J. Li, and J. Wang, "Covert communication achieved by a greedy relay in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4766–4779, Jul. 2018.

[37] K. Shahzad, X. Zhou, S. Yan, J. Hu, F. Shu, and J. Li, "Achieving covert wireless communications using a full-duplex receiver," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8517–8530, Dec. 2018.

[38] J. Wang, W. Tang, Q. Zhu, X. Li, H. Rao, and S. Li, "Covert communication with the help of relay and channel uncertainty," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 317–320, Feb. 2019.

[39] B. He, S. Yan, X. Zhou, and V. K. N. Lau, "On covert communication with noise uncertainty," *IEEE Commun. Lett.*, vol. 21, no. 4, pp. 941–944, Apr. 2017.

[40] R. Soltani, D. Goeckel, D. Towsley, B. A. Bash, and S. Guha, "Covert wireless communication with artificial noise generation," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7252–7267, Nov. 2018.

[41] M. R. Bloch, "Covert communication over noisy channels: A resolvability perspective," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2334–2354, May 2016.

[42] R. W. Heath, Jr., N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2017.

[43] Q. Zhang, S. Jin, K.-K. Wong, H. Zhu, and M. Matthaiou, "Power scaling of uplink massive MIMO systems with arbitrary-rank channel means," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 966–981, Oct. 2014.

XINGBO LU received the B.S. degree from the Beijing Institute of Technology, in 2017. He is currently pursuing the Ph.D. degree in communications and information system with the Institute of Communications Engineering, Army Engineering University of PLA. His research interests include physical layer security, relaying networks, millimeter wave communication, and low probability of detection communication.

WEIWEI YANG (S'08–M'11) received the B.Sc., M.Sc., and Ph.D. degrees in telecommunications from the PLA University of Science and Technology, Nanjing, China, in 2003, 2006, and 2011, respectively. He is currently an Associate Professor with the College of Communication Engineering, Army Engineering University of PLA. He is also a coauthor of the book *Handbook of Cognitive Radio* (Springer, 2017). His research interests include cooperative communications, cognitive radio, and physical layer security. He has served as a TPC Member for WCSP 2011/2014/2017/2018, GC 2016 Workshops, GC 2017 Workshops, and ICC 2016-Workshops. He was a co-recipient of the Best Paper Award from WCSP 2011. He has served as the Publication Co-Chair for WCSP 2015 and the Track Chair for IEEE CIC ICCC 2017 and WCSP 2019.

YUEMING CAI (M'05–SM'12) received the B.S. degree in physics from Xiamen University, Xiamen, China, in 1982, and the M.S. degree in microelectronics engineering and the Ph.D. degree in communications and information systems from Southeast University, Nanjing, China, in 1988 and 1996, respectively. His current research interests include MIMO systems, OFDM systems, signal processing in communications, cooperative communications, and wireless sensor networks.

XINRONG GUAN received the B.S. degree in communications engineering and the Ph.D. degree in communications and information systems from the Institute of Communications Engineering, PLA University of Science and Technology, Nanjing, China, in 2009 and 2014, respectively. His current research interests include physical layer security, wireless key generation, cooperative communications, and cognitive radio networks.