This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2841027, IEEE Access

Performance Analysis of Secret Precoding-Aided Spatial Modulation with Finite-Alphabet Signaling

Feilong Wu, Wenjie Wang, *Member, IEEE*, Chen Dong, Lie-Liang Yang, *Fellow, IEEE*

***Abstract*—The precoding-aided spatial modulation (PSM) is firstly characterized with respect to its secrecy performance, which shows that PSM is a secrecy embedded communication scheme when operated in time-division duplex (TDD) mode, but experiences security risk under frequency-division duplex (FDD) mode. Therefore, we extend the PSM to a secret PSM (SPSM) that is suitable for operation in any communications scenarios experiencing passive eavesdropping. In the proposed SPSM, the secrecy performance is enhanced by using a part of transmit power to interfere the eavesdropper's reception. In this paper, the secrecy performance of SPSM with discrete inputs is comprehensively investigated in terms of both bit error rate (BER) and secrecy rate, and by both analysis and simulations. Relying on the asymptotic analysis, we also derive the upper- and low-bounds for the error performance and secrecy rate of SPSM. Furthermore, we study the power-allocation between information and interference transmission, in order to maximize SPSM's secrecy performance. Our studies result in a range of expressions, which are validated by Monte-Carlo simulations. Furthermore, our studies demonstrate that the analytical formulas are beneficial to the optimization of SPSM systems for maximizing its security performance.**

***Index Terms*—Spatial modulation (SM), bit error rate (BER), secrecy capacity, power-allocation, artificial noise, finite-alphabet inputs.**

I. INTRODUCTION

WIRELESS communications may suffer from eavesdropping attack due to its broadcast nature. Spread-spectrum techniques [1] have usually been used to resist eavesdropping for achieving low-probability-interception (LPI) transmission. However, their security relies on the privacy of the pseudo-noise (PN) sequences shared between a transmitter and its desired receiver. To circumvent this shortcoming, precoding-aided spatial modulation (PSM) [2–4] and directional modulation (DM) [5–8] have been redesigned for secret wireless transmission without requiring pre-shared PN sequences or secret keys between a transmitter-receiver pair. In this paper, our focus is on the PSM-assisted secrecy communications. In PSM systems, there are two types of modulations employed for information transmission over MIMO channels, which are the space shift keying (SSK) [9] and conventional amplitude-phase modulation (APM). The SSK is implemented using the indices of receiver antennas, with the aid of a zero-forcing precoder (ZFP) or a minimum mean square error precoder (MMSEP). PSM employs two distinctive advantages [10, 11]: i) information transmission in spatial domain, and ii) low-complexity detection provided by the embedded precoding. Owing to its merits, PSM has been regarded as a promising signaling scheme for operation in conjunction with massive MIMO, millimeter wave (mm-Wave), etc., of the 5th generation (5G) wireless techniques [12–15]. In [3, 4], we have identified PSM's third distinctive advantage for facilitating secret communications. In PSM, the portion of information conveyed by the SSK acts actually as a PN sequence, which randomly activates one of the receive antennas to enable the receive antenna's activation patterns to be controlled directly by the random source information. At the receiver side, this information can be readily decoded by observing the antenna activation patterns. However, an eavesdropper will be much hard to demodulate the information carried by the SSK of the PSM. This is because the precoder, such as ZFP, MMSEP, etc., is designed only based on the channel state information (CSI) of the desired receiver and, hence, the antenna activation patterns do not appear at the eavesdropper. Therefore, PSM can be a secret, or at least, a LPI transmission scheme, when it is operated in the scenarios where eavesdroppers are hard to estimate the channels between a transmitter-receiver pair, and consequently, hard to know the precoder in the ZFP or MMSEP principles.

The above arguments hold true, when the PSM is operated in the time-division duplex (TDD) mode, where no CSI feedback from receiver to transmitter is required. However, when operated in the frequency-division duplex (FDD) mode, a feedback link may be required to send the CSI estimated at a receiver to its transmitter. By overhearing this feedback channel, an eavesdropper may obtain the CSI sent by the receiver and hence, knows the precoder. In this case, the eavesdropper may also reliably detect the information sent by the PSM, provided that it has a sufficient signal-to-noise ratio (SNR) for detection, as shown later in Section II. In order to prevent from eavesdropping in both TDD and FDD modes, we need to enhance the PSM by designing a so-called secret precoder, which should not be recovered by any eavesdropper, even when it knows the exact CSI used for designing the precoder. With this motivation, in this paper, we propose a secret PSM (SPSM), which operates a time-varying precoder obtained by adding some artificial perturbation (AP) to the ZFP. As shown in our forthcoming discourses, our SPSM

Feilong Wu is with China Academy of Space Technology, Xi'an Branch (CAST Xi'an), Xi'an, 710100, Shaanxi, China (e-mail: wufeilong@stu.xjtu.edu.cn). This work was completed when he was a PhD student of Xi'an Jiaotong University. Wenjie Wang is with the Ministry of Education Key Lab for Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an, 710049, Shaanxi, China (e-mail: wjwang@xjtu.edu.cn). Chen Dong is with the Huawei Device Co., Ltd., China, (e-mail: dongchen403@gmail.com), and Lie-Liang Yang is with the School of Electronics and Computer Science, University of Southampton, SO17 1BJ, UK (e-mail: lly@ecs.soton.ac.uk). Their work was supported in part by the NSFC (NO. 61671366) of China and the EPSRC (No. EP/P034284/1) of the UK.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2841027, IEEE Access

employs all the advantages of the original PSM, such as, low-complexity detection, when viewed from the desired receiver. By contrast, from an eavesdropper, the time-variant AP added to the precoder makes its detection much hard.

A. Related Work

In literature, there are several references, e.g. [16–20], having investigated the security issues of the conventional SM [21], which uses the indices of transmit antennas to convey information in the spatial domain. Since the conventional SM usually uses only one to several radio-frequency (RF) chains for transmission, the multiple transmit antennas may not be effectively exploited for preventing from eavesdropping. By contrast, the PSM with MIMO precoding employs the capability to elevate the secrecy performance of wireless transmission, in addition to its simplified receiver [3, 4, 22]. For example, PSM's secrecy performance can be significantly improved by designing a precoding matrix so as to maximize the ratio between the received power of the desired receiver and that of the eavesdropper [3]. However, this optimization requires the transmitter to know the eavesdropper's CSI. Hence, it is only suitable for the communication scenarios where an active eavesdropper exists, but not suitable for the communication environments conflicting a passive eavesdropper. In [4], we have firstly extended the PSM to the SPSM for mitigating passive eavesdropping, and obtained the optimum power-allocation between ZFP and AP. However, the power-allocation in [4] has to be dependent on the Monte Carlo simulations, which imposes a heavy computation burden on the transmitter, especially, when the system has a large input-output dimension or uses a higher order of APM.

The power-allocation in SPSM systems shares the similarity as that in the MIMO wiretap channels, where artificial noise (AN) [23] is used to cripple the reception of passive eavesdroppers. Therefore, below we review some recent progress in AN's power-allocation. In [24], the authors have first considered the optimal power-allocation between an information-bearing signal and an AN, when assuming communication over multiple-input single-output (MISO) Rayleigh fading channels. Specifically, under the assumption of a noiseless eavesdropper, a closed-form formula for the lower bound of the achievable secrecy rate has first been derived, which is then used to optimize the power allocation. In more detail, let the numbers of antennas employed at the transmitter, receiver and the eavesdropper be N_t, N_r and N_e, respectively. The studies in [24] show that when operated in high SNR region, it is near-optimal to distribute half of the transmission power to the AN in the case of conflicting a MISO single-antenna eavesdropper (MISOSE), and $\sqrt{N_e}/(1+\sqrt{N_e})$ of the transmission power to the AN in the case of MISO multi-antenna eavesdropper (MISOME). In [25], the same assumption of noiseless eavesdropper has been considered for the case of MIMO multi-antenna eavesdropper (MIMOME). However, the results in [24, 25] have been obtained by considering the worst case and also ignoring thermal noise at the eavesdropper, making them usually not suitable for application in low SNR region. In [26], tight bounds for the ergodic secrecy rate of various MIMO systems using AN have been studied, when assuming that there are a large number of transmit antennas. From the studies in [26] we are implied that $N_e/(N_r+N_e)$ of the transmission power should be allocated to the AN in high SNR region, while no power is allocated to the AN in low SNR region.

All the above-mentioned references have assumed the Gaussian input signals for deriving the closed-form expressions for the ergodic secrecy rate. Despite the information theoretic optimality of the Gaussian input signals, practical signals for communications are typically non-Gaussian, and constructed from a finite alphabet, such as, phase-shift keying (PSK) symbols, quadrature-amplitude modulation (QAM) symbols, etc. In this case, power-allocation in physical-layer security systems becomes more challenging. This is because the MIMO channels with finite-alphabet inputs belong to the category of the discrete-input continuous-output memoryless channels (DCMC) [27, 28], whose closed-form expressions for rate/capacity are extremely hard to obtain. Consequently, Monte Carlo simulations are usually relied on for finding the DCMC's capacity [27–30]. Due to the absence of closed-form formulas for the secrecy rate of DCMC channels, the relationship existing between the mutual information of MIMO systems and the MMSE detector may be exploited for transmit power-control [31] or precoder optimization [32]. However, to the best of our knowledge, there is short of research about the optimization of AN, including both the AN signaling design and power-allocation, in the context of the DCMC channels. Our SPSM studied in this contribution belongs to this category, and there are no closed-form expressions available from references for its secrecy rate. Hence, as previous mentioned, the optimization for power-allocation between ZFP and AP is highly challenging.

B. Contributions and Outcomes

In this paper, we first demonstrate the principles of the SPSM and, then, investigate the secrecy performance of the SPSM, which includes both the bit error rate (BER) performance and the secrecy capacity. Our studies show that in the general cases, it is extremely hard (if it is not impossible) to derive the meaningful closed-form expressions for the error and secrecy performance, based on which we can derive the optimum power-allocation. For this sake, we carry out the asymptotic analysis by assuming that the SPSM system has a big number of transmit antennas, which is significantly larger than the number of antennas at the receiver and also that at the eavesdropper.

Specifically, when the error performance is considered, if the SPSM system uses a big number of transmit antennas, the artificial interference presenting at the eavesdropper can be approximated as additive white Gaussian noise. In this case, the equivalent signal to interference plus noise ratio (SINR) at the eavesdropper is determined by the power assigned by the transmitter to the AP. Therefore, the transmitter can directly control the error performance of the eavesdropper's detection via assigning a corresponding power to the AP. Furthermore, under the asymptotic assumption, we can derive the BER union-bounds for both the receiver and the eavesdropper.

Based on the above observations and with the aid of the BER union-bounds, the transmitter of SPSM is capable of carrying out the corresponding power-allocation, so as to make the error performance at the eavesdropper worst, while maintaining the error performance at the receiver above the required level.

In the context of the secrecy capacity, we first provide an exact expression for the secrecy rate, which is however not in closed-form. Based on the formula obtained, the secrecy capacity can be obtained by maximizing the achievable secrecy rate via optimum power-allocation. However, in this case, the secrecy rate has to be evaluated by Monte Carlo simulations, and the secrecy capacity has to be obtained via the exhaustive search by using a huge number of realizations of the possible power-allocation. By contrast, under the asymptotic scenario where the number of transmit antennas is large, we are able to derive a closed-form formula for the approximate secrecy rate. From this formula we can readily find a near-optimal power-allocation solution via the one dimensional search, which is not dependent on the complicated computations as Monte-Carlo simulation. Our studies show that, although our analytical results are obtained in the asymptotic cases and by imposing some approximation, the power-allocation obtained in this way is in general close to the optimum one, even when the number of transmit antennas is limited, as demonstrated by our numerical and simulation results shown in Section VI. In addition to the above-mentioned, in this paper, we furthermore derive the asymptotic results for achieving the near-optimal power-allocation, when the SPSM is operated in either low or high SNR region.

The rest of this paper is structured as follows. We begin with considering the PSM over the MIMO wiretap channels in section II, where we also explain the LPI characteristics of the PSM under the TDD mode. In section III, the secrecy performance of the PSM is enhanced by introducing the SPSM, which is capable of providing secure transmission under both the TDD and FDD modes. In the following two sections, i.e., in section IV and section V, we analyze both the error performance and the capacity performance of the SPSM, based on which we address the power-allocation issues. The performance results and discussion are provided in section VI and, finally, we conclude this paper in section VII.

Notation: Boldface uppercase and lowercase variables denote matrices and vectors, respectively. $\boldsymbol{A}^H$, $\boldsymbol{A}^{-1}$, $tr(\boldsymbol{A})$ represent respectively the Hermitian transpose, inverse, and trace operation of a matrix $\boldsymbol{A}$. $\boldsymbol{I}_N$ stands for a N-dimensional identity matrix. $\mathbb{C}^{m\times n}$ stands for the complex space of $m\times n$ dimensions. $\|\cdot\|$ represents the Euclidean norm of a vector; $|\cdot|$ and $\Re\{\cdot\}$ take the modulus and real part of a complex number, respectively. $(x)^+=\max(0,x)$ and $\mathbb{E}[\cdot]$ evaluates the expectation with respect to all the random variables within the bracket. Finally, $\log$ is used for the base two logarithm while $Q(x)$ is the Gaussian Q-function defined as $Q(x)=\frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-t^2/2}dt$.

II. PSM over MIMOME Channels

We consider the communication scenario that a transmitter sends messages to a receiver in the presence of a passive eavesdropper. The number of antennas of the transmitter, receiver and eavesdropper are denoted as N_t, N_r and N_e, respectively. Following the principles of PSM [2, 10], we assume that $N_t>N_r,N_e$, and furthermore, $N_r=2^{k_1}$ with k_1 being a positive integer. In the considered PSM, only one of the receiver's antennas is activated and its index i $(i=1,\cdots,N_r)$ is used to convey k_1 bits per symbol. This modulation scheme is referred to as the SSK for convenience. Besides the SSK, the transmitter also employs a conventional $M(=2^{k_2})$-ary APM, such as M-ary quadrature amplitude modulation (MQAM) or M-ary phase-shift keying (MPSK), to send another k_2 bits per symbol to the receiver. Let us denote the APM constellation set as $\mathcal{B}=\{b_1,\cdots,b_M\}$. Then, by combining the SSK symbol and APM symbol, a PSM symbol can be represented as

$$\boldsymbol{s}_i^j=\boldsymbol{e}_ib_j, \tag{1}$$

where $\boldsymbol{e}_i$ is the i-th column of an identity matrix $\boldsymbol{I}_{N_r}$, representing that the ith antenna of the receiver is activated, while $b_j\in\mathcal{B}$ and satisfies $\mathbb{E}[|b_j|^2]=1$. Hence, each PSM symbol can convey in total $k=k_1+k_2$ bits information.

Let $\boldsymbol{P}=[\boldsymbol{p}_1,\cdots,\boldsymbol{p}_{N_r}]\in\mathbb{C}^{N_t\times N_r}$ be the preprocessing matrix used by the transmitter to process PSM symbol. Then, the signals received by the receiver and eavesdropper can be respectively expressed as

$$\boldsymbol{y}=\boldsymbol{H}\boldsymbol{P}\boldsymbol{s}_i^j+\boldsymbol{u}, \tag{2}$$

$$\boldsymbol{z}=\boldsymbol{H}_e\boldsymbol{P}\boldsymbol{s}_i^j+\boldsymbol{v}, \tag{3}$$

where $\boldsymbol{u}\sim\mathcal{CN}(\boldsymbol{0},\sigma_u^2\boldsymbol{I}_{N_r})$ and $\boldsymbol{v}\sim\mathcal{CN}(\boldsymbol{0},\sigma_v^2\boldsymbol{I}_{N_e})$ are the additive white Gaussian noise (AWGN). As in [26, 31, 32], for simplicity, we assume the same noise power at both the receiver and the eavesdropper, i.e., $\sigma_u^2=\sigma_v^2=1/\gamma$, where γ is the average SNR per symbol. In (2) and (3), the matrices $\boldsymbol{H}\in\mathbb{C}^{N_r\times N_t}$ and $\boldsymbol{H}_e\in\mathbb{C}^{N_e\times N_t}$ represent the channels from the transmitter to the receiver and that from the transmitter to the eavesdropper, respectively. We assume that the transmitter has full knowledge of $\boldsymbol{H}$ via the receiver's feedback channels or using channels' reciprocity, but no knowledge about $\boldsymbol{H}_e$. The eavesdropper is assumed to be able to estimate its own channel $\boldsymbol{H}_e$, using for example, the pilots sent by the transmitter. However, whether the eavesdropper is able to acquire the receiver's channel $\boldsymbol{H}$ is depended on the duplex mode operated between the transmitter and receiver, which will be analyzed at the end of this section. Additionally, all the channels are assumed to experience block Rayleigh fading.

The original PSM [2] has been motivated to achieve low-complexity detection by designing a precoder $\boldsymbol{P}$, which, according to the information transmitted, uses beamforming to activate only one of the receive antennas, while keeps the other receive antennas silent. This can be implemented with the aid of the ZFP or MMSEP. For the purpose to show the principles, below we consider only the ZFP, which uses the preprocessing matrix of

$$\boldsymbol{P}=\beta\boldsymbol{H}^H(\boldsymbol{H}\boldsymbol{H}^H)^{-1}, \tag{4}$$

where $\beta=\sqrt{N_r/tr[(\boldsymbol{H}\boldsymbol{H}^H)^{-1}]}$ is a normalization factor for achieving the power constraint of $tr(\boldsymbol{P}\boldsymbol{P}^H)=N_r$. The ZFP is capable of completely canceling the inter-antenna interference at the receiver, which can be seen by substituting (4) into (2), yielding

$$\boldsymbol{y} = \beta \boldsymbol{s}_i^j + \boldsymbol{u}. \tag{5}$$

Furthermore, it can be shown that (5) can be decomposed into

$$y_i = \beta b_j + u_i; \quad y_k = u_k, \quad \forall k \neq i. \tag{6}$$

From (6) we can know that information only appears on the ith antenna of the receiver, when assuming that a k_1-bit SSK symbol having a value of i was transmitted. All the other antennas of the receiver output only noise.

As shown in [2], the k_1-bit SSK symbol and the k_2-bit APM symbol can be detected using either joint detection (JD) or successive detection (SD). When the JD is employed, the maximum likelihood (ML) detector gives

$$\begin{aligned}\left\langle \hat{i}, \hat{b}_j \right\rangle &= \arg \min_{i \in [1, N_r], b_j \in \mathcal{B}} ||\boldsymbol{y} - \beta \boldsymbol{s}_i^j||^2 \\ &= \arg \min_{i \in [1, N_r], b_j \in \mathcal{B}} \left[\beta^2 |b_j|^2 - 2\Re\{y_i^* b_j\} \right]. \end{aligned} \tag{7}$$

By contrast, when the SD is employed, the APM symbol is detected after the detection of the SSK symbol as

$$\hat{\imath} = \arg \max_{i \in [1, N_r]} |y_i|^2, \quad \hat{b}_j = \arg \min_{b_j \in \mathcal{B}} |y_{\hat{\imath}}/\beta - b_j|^2. \tag{8}$$

In comparison with the JD, the SD can significantly reduce the detection complexity, but at the cost of marginal degradation of the achievable BER performance [10]. From the above-described detection principles we can know that, no matter is JD or SD employed, the receiver is capable of achieving the detection, provided that it knows β, instead of the CSI of $\boldsymbol{H}$. Furthermore, when the number of transmit antennas is large, β converges to a constant of $1/\sqrt{N_t}$. In this case, the desired receiver in fact carries out non-coherent detection, which facilitates practical implementation.

Let us now consider how the eavesdropper may intercept the private messages conveyed from the transmitter to the receiver using PSM. Since the precoder $\boldsymbol{P}$ is designed based only on the CSI from the transmitter to the receiver, inter-antenna interference still exists among the antennas of the eavesdropper. Therefore, the eavesdropper is unable to demodulate the PSM signal using the low-complexity detection like (7) or (8), even if it employs perfect $\boldsymbol{H}_e$. Explicitly, (2) and (3) can also be represented in the forms of

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{p}_i b_j + \boldsymbol{u}, \tag{9}$$

$$\boldsymbol{z} = \boldsymbol{H}_e \boldsymbol{p}_i b_j + \boldsymbol{v}. \tag{10}$$

From these equations we have the following observations. Firstly, the PSM acts as a precoding vector hopper, with the hopping pattern $\{\boldsymbol{p}_i\}$ formed from $\boldsymbol{P}$ under the control of the random messages sent from the transmitter to the receiver. If the channels from the transmitter to the receiver vary sufficiently fast, the transmission from the transmitter to the receiver appears more or less as an 'one-time pad' cryptographic scheme, which is rendered absolutely secure. Secondly, owing to the ZFP, and if the SNR at the receiver is sufficiently high, the receiver can readily capture the hopping pattern via observing its activated antennas. In this way, the receiver can fulfill the detection based on (7) or (8). Thirdly, the situations for the eavesdropper are very different. As above-mentioned, when the channels from the transmitter to the receiver vary fast, the eavesdropper is incapable of detecting any information sent from the transmitter to the receiver due to the one-time pad effect. By contrast, if the channels from the transmitter to the receiver vary slow, so that the eavesdropper is capable of estimating the joint of $\boldsymbol{H}_e\boldsymbol{P}$, as seen in (3), the eavesdropper may employ a blind detection scheme [35, 36] to detect the APM symbol b_j. However, the eavesdropper is still incapable of detecting the SSK symbol i, as it is unable to estimate a separate $\boldsymbol{P}$. In order to detect both i and b_j, the eavesdropper has to know both its CSI $\boldsymbol{H}_e$ and the perfect knowledge of $\boldsymbol{P}$ (or $\boldsymbol{H}$). If these are available, the eavesdropper can then implement the JD as

$$\left\langle \hat{i}, \hat{b}_j \right\rangle = \arg \min_{\boldsymbol{p}_i \in \{\boldsymbol{p}_1, \cdots, \boldsymbol{p}_{N_r}\}, b_j \in \mathcal{B}} \|\boldsymbol{z} - \boldsymbol{H}_e \boldsymbol{p}_i b_j\|^2. \tag{11}$$

Note that, in TDD systems, the transmitter can estimate the downlink channels $\boldsymbol{H}$ using the uplink training signals transmitted by the receiver based on the channels' reciprocity. Since the transmitter and the eavesdropper are separated in space, $\boldsymbol{H}$ and $\boldsymbol{H}_e$ should usually be independent. Consequently, the eavesdropper hardly knows $\boldsymbol{H}$ and, hence, $\boldsymbol{P}$, even though it can overhear the pilot signals sent from the receiver. Without the knowledge of $\boldsymbol{P}$, the eavesdropper is unable to implement the detection in the way as shown (11). Therefore, when operated in the TDD systems, the original PSM as a whole belongs to a LPI communication scheme, which can provide confidential communication between the transmitter and receiver. By contrast, in FDD systems, the transmitter may have to acquire $\boldsymbol{H}$ from the receiver via feedback channels. In these communication systems, the eavesdropper may have the chances to know $\boldsymbol{P}$ by overhearing $\boldsymbol{H}$ sent over the feedback channels from the receiver to the transmitter. Then, the eavesdropper may decode the messages using (11). In other words, the PSM is not robust to the eavesdropping attack in FDD systems. In the following discourses, we propose a secret PSM (SPSM) scheme that is efficient for operation in both TDD and FDD modes.

III. Secret Precoding-Aided Spatial Modulation

From the analysis in Section II, we can know that the knowledge of $\boldsymbol{P}$ to the eavesdropper determines the security level of the PSM system. When the transmitter-receiver channels vary slow, the precoder $\boldsymbol{P}$ in (4) also varies slow, which is beneficial to eavesdropping. Therefore, the basic principles behind our SPSM is to construct a precoder $\boldsymbol{P}$, which is capable of retaining the ZFP nature at the desired receiver, while presenting fast time-varying at the eavesdropper, regardless of the time-varying rate of the transmitter-receiver channels. In this way, the one-time pad effect is always experienced by the eavesdropper and, hence, the secrecy performance of the SPSM can be significantly improved.

Let us apply the singular value decomposition on $\boldsymbol{H}$, yielding

$$\boldsymbol{H} = \boldsymbol{U}[\boldsymbol{\Lambda} \quad \boldsymbol{0}][\boldsymbol{V}_1 \quad \boldsymbol{V}_0]^H, \tag{12}$$

where $\boldsymbol{U} \in \mathbb{C}^{N_r \times N_r}$ is a unitary matrix, $\boldsymbol{\Lambda}$ is a $(N_r \times N_r)$-dimensional diagonal matrix containing the non-zero singular values, $\mathbf{0}$ is an all-zero matrix, $\boldsymbol{V}_1 \in \mathbb{C}^{N_t \times N_r}$ and $\boldsymbol{V}_0 \in \mathbb{C}^{N_t \times (N_t - N_r)}$ determine respectively a signal space and a null space for $\boldsymbol{H}$. Based on (12), the time-varying precoder for the SPSM can be structured as

$$\begin{aligned} \boldsymbol{P} &= \bar{\boldsymbol{P}} + \tilde{\boldsymbol{P}} \\ &= \zeta \boldsymbol{H}^H (\boldsymbol{H}\boldsymbol{H}^H)^{-1} + \boldsymbol{V}_0 \boldsymbol{R} \end{aligned} \tag{13}$$

where ζ is a coefficient that can be adjusted to distribute the transmit power between $\bar{\boldsymbol{P}}$ and $\tilde{\boldsymbol{P}}$, while $\boldsymbol{R} = [\boldsymbol{r}_1, \cdots, \boldsymbol{r}_{N_r}] \in \mathbb{C}^{(N_t - N_r) \times N_r}$ is a random matrix whose entries are complex Gaussian random variables with zero mean and a variance σ_r^2.

As seen in (13), the precoder of the SPSM contains two components: a basic ZFP of $\bar{\boldsymbol{P}}$ and a time-varying AP of $\tilde{\boldsymbol{P}}$. Since $\tilde{\boldsymbol{P}}$ is constructed from $\boldsymbol{V}_0$ and lies in the null-space of $\boldsymbol{H}$, we have $\boldsymbol{H}\tilde{\boldsymbol{P}} = \mathbf{0}$. Hence, the AP is automatically canceled, when the transmitted signals arrive at the receiver. In detail, when substituting (13) into (2) and (3), we obtain

$$\boldsymbol{y} = \zeta \boldsymbol{s}_i^j + \boldsymbol{u}, \tag{14}$$

$$\boldsymbol{z} = \boldsymbol{H}_e \bar{\boldsymbol{P}} \boldsymbol{s}_i^j + \boldsymbol{H}_e \boldsymbol{V}_0 \boldsymbol{R} \boldsymbol{s}_i^j + \boldsymbol{v}. \tag{15}$$

Eq. (14) shows that, for the desired receiver, the SPSM inherits all the advantages of the PSM, including the low-complexity detection of (8). By contrast, for the eavesdropper, attacking the SPSM becomes much harder than attacking the PSM because of the additional interference $\boldsymbol{V}_0 \boldsymbol{R} \boldsymbol{s}_i^j$, as seen in (15). Furthermore, considering the statistics of the random vectors $\boldsymbol{R}\boldsymbol{s}_i^j = \boldsymbol{r}_i b_j$, if b_j are MPSK symbols of unit power, we have $\boldsymbol{r}_i b_j \sim \mathcal{CN}(\mathbf{0}, \sigma_r^2 \boldsymbol{I}_{N_r})$. By contrast, if b_j are QAM symbols with unit average power, $\boldsymbol{r}_i b_j$ may be approximated as the Gaussian random variable obeying the distribution of $\mathcal{CN}(\mathbf{0}, \sigma_r^2 \boldsymbol{I}_{N_r})$. Hence, in principle, the AP in our proposed SPSM acts similarly as the well-known AN [23]. In the SPSM, AP can protect not only the APM symbols but also the SSK symbols, which will become explicit in the forthcoming analysis.

Let the equivalent noise in (15) seen by the eavesdropper be denoted as $\boldsymbol{w} = \boldsymbol{H}_e \boldsymbol{V}_0 \boldsymbol{R} \boldsymbol{s}_i^j + \boldsymbol{v}$, whose covariance matrix is given by

$$\boldsymbol{\Sigma} = \mathbb{E}[\boldsymbol{w}\boldsymbol{w}^H] = \sigma_r^2 \boldsymbol{H}_e \boldsymbol{V}_0 \boldsymbol{V}_0^H \boldsymbol{H}_e^H + \sigma_v^2 \boldsymbol{I}. \tag{16}$$

Under the worst scenario where the eavesdropper knows ideally $\boldsymbol{H}$, and hence it can derive $\bar{\boldsymbol{P}}$, the eavesdropper can then carry out the optimal detection based on (15) as [40]

$$\left\langle \hat{i}, \hat{b}_j \right\rangle = \arg \min_{i \in [1, N_r], b_j \in \mathcal{B}} \|\boldsymbol{\Sigma}^{-\frac{1}{2}} \left(\boldsymbol{z} - \boldsymbol{H}_e \bar{\boldsymbol{P}} \boldsymbol{s}_i^j \right) \|^2. \tag{17}$$

Explicitly, the above detection performance of (17) is related to σ_r^2. Hence according to (13), there exists the power-allocation between $\bar{\boldsymbol{P}}$ and $\tilde{\boldsymbol{P}}$, which yields a trade-off between the reliability and security of the transmitter-receiver link. Let the fraction of the total transmit power allocated to $\bar{\boldsymbol{P}}$ be denoted as θ $(0 < \theta \leq 1)$. Then, using the power-constraints $tr(\bar{\boldsymbol{P}}\bar{\boldsymbol{P}}^H) = \theta N_r$ and $\mathbb{E}\left[tr(\tilde{\boldsymbol{P}}\tilde{\boldsymbol{P}}^H)\right] = (1 - \theta) N_r$, we can obtain

$$\zeta = \sqrt{\frac{\theta N_r}{tr[(\boldsymbol{H}\boldsymbol{H}^H)^{-1}]}}, \quad \sigma_r^2 = \frac{1 - \theta}{N_t - N_r}. \tag{18}$$

From (18) we know that the PSM constitutes a special case of the SPSM with $\theta = 1$. Furthermore, we can be implied from (18) that the value of θ imposes significant impact on the achievable secrecy performance. In the following sections, two metrics, namely, error probability and secrecy rate, are considered for studying the achievable performance of the SPSM scheme. More specifically, in Section IV, we study the BER performance of the SPSM, when the value of θ is fixed for all the considered SNRs. By contrast, in Section V, we aim at maximizing the secrecy rate of the SPSM by finding the optimal value of θ for every given SNR. As a token of the efforts, our analytical results are beneficial for a SPSM transmitter to choose a desirable value of θ, in order to maximize its secrecy performance.

IV. Analysis of Error Probability

In this section, we analyze the error performance of both the receiver and eavesdropper. As seen in (15), in order to analyze the error probability of the eavesdropper, we need to know the statistics of the covariance matrix $\boldsymbol{\Sigma}$ of (16). Explicitly, the covariance matrix $\boldsymbol{\Sigma}$ is not a diagonal matrix, and analyzing the performance of (17) in general requires the random matrix theory and is very complicated. However, when N_t is large resulting in a massive MIMO [37] systems, the covariance matrix $\boldsymbol{\Sigma}$ converges to a diagonal matrix and in this case, the performance analysis of SPSM systems becomes feasible. Therefore, in this section, we analyze the asymptotic performance of the SPSM systems, when assuming that the transmitter, which can be a base station, employs a large number of antennas, i.e., $N_t \gg N_r, N_e$. In this case, by applying (18), the covariance matrix of (16) can be simplified to

$$\begin{aligned} \boldsymbol{\Sigma} =& \lim_{N_t \to \infty} \frac{1 - \theta}{N_t - N_r} \boldsymbol{H}_e \boldsymbol{V}_0 \boldsymbol{V}_0^H \boldsymbol{H}_e^H + \sigma_v^2 \boldsymbol{I} \\ =& [(1 - \theta) + \sigma_v^2] \boldsymbol{I}. \end{aligned} \tag{19}$$

From this we can know that $\boldsymbol{w}$ converges to a white Gaussian noise vector, which has zero mean and a covariance matrix $[(1 - \theta) + \sigma_v^2] \boldsymbol{I}_{N_e}$. Consequently, the detector of (17) can be simplified to

$$\left\langle \hat{i}, \hat{b}_j \right\rangle = \arg \min_{i \in [1, N_r], b_j \in \mathcal{B}} \|\boldsymbol{z} - \boldsymbol{H}_e \bar{\boldsymbol{P}} \boldsymbol{s}_i^j \|^2, \tag{20}$$

whose detection performance is a function of the average SINR at the eavesdropper, which can be derived from (19) as

$$\gamma_e = \frac{\gamma \theta}{\gamma (1 - \theta) + 1}. \tag{21}$$

From (21) we can see that $\gamma_e \approx \theta / (1 - \theta)$, when $\gamma \to \infty$. This explains that, when the system's SNR is sufficiently high, the eavesdropper's BER does not decrease with the increase of γ, but exhibits an error floor determined by θ, which is controllable by the transmitter using power allocation.

Below we derive the tight analytical BER bounds of both the receiver using the detector of (7) and the eavesdropper using the detector of (20), when assuming that all the channels experience Rayleigh fading. From our analysis and performance results, we can gain the insight into the power-allocation between ZFP and AP by adjusting the value of θ.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2841027, IEEE Access

A. Bit Error Rate of the Receiver

Let us define a set $\mathcal{S}$ with cardinality MN_r, which contains all the possible PSM symbols. In some cases when there is no confusion, we denote the transmitted PSM symbol $\boldsymbol{s}_i^j$ as $\boldsymbol{s}_{\boldsymbol{\tau}}$, while another PSM symbol $\boldsymbol{s}_m^n$ as $\boldsymbol{s}_\varepsilon$, in order to make the derivations more readable.

Following the principles of digital communications, the union bound for the receiver's BER can be expressed [38]

$$\begin{aligned} P_e(R) \leq & \frac{1}{kMN_r} \sum_{\boldsymbol{s}_i^j \in \mathcal{S}} \sum_{\boldsymbol{s}_m^n \in \mathcal{S}, \boldsymbol{s}_m^n \neq \boldsymbol{s}_i^j} e(\boldsymbol{s}_i^j, \boldsymbol{s}_m^n) \\ & \times \mathbb{E}_{\boldsymbol{H}} \left\{ P\left(\boldsymbol{s}_i^j \mapsto \boldsymbol{s}_m^n \,|\boldsymbol{H} \right) \right\} \end{aligned} \tag{22}$$

where $P(\boldsymbol{s}_i^j \mapsto \boldsymbol{s}_m^n \,|\boldsymbol{H}\,)$ denotes the pairwise error probability (PEP) between two PSM symbols $\boldsymbol{s}_i^j$ and $\boldsymbol{s}_m^n$ under a given channel realization $\boldsymbol{H}$, while $e(\boldsymbol{s}_i^j, \boldsymbol{s}_m^n)$ is the number of different bits between $\boldsymbol{s}_i^j$ and $\boldsymbol{s}_m^n$. According to the decision variable given in (14), when the JD is used, the conditional PEP in (22) is equal to

$$\begin{aligned} P(\boldsymbol{s}_i^j \mapsto \boldsymbol{s}_m^n \,|\boldsymbol{H}\,) &= P\left\{ \|\boldsymbol{y} - \zeta \boldsymbol{s}_i^j\|^2 > \|\boldsymbol{y} - \zeta \boldsymbol{s}_m^n\|^2 \right\} \\ &= P\left\{ \Re\left[\boldsymbol{u}^H \left(\boldsymbol{s}_m^n - \boldsymbol{s}_i^j \right) \right] > \frac{1}{2}\zeta \left\| \boldsymbol{s}_m^n - \boldsymbol{s}_i^j \right\|^2 \right\}. \end{aligned} \tag{23}$$

In the above formulas, we can show that $\Re\left[\boldsymbol{u}^H \left(\boldsymbol{s}_m^n - \boldsymbol{s}_i^j \right) \right]$ is a real random variable obeying the distribution of $\mathcal{N}(0, \frac{\sigma_u^2}{2} \|\boldsymbol{s}_m^n - \boldsymbol{s}_i^j\|^2)$. Hence, using the Gaussian Q-function [38], we can obtain

$$P\left(\boldsymbol{s}_i^j \mapsto \boldsymbol{s}_m^n \,|\boldsymbol{H} \right) = Q\left(\zeta \sqrt{\frac{d_{\tau\varepsilon}}{2\sigma_u^2}} \right) \tag{24}$$

where $d_{\tau\varepsilon}$ is defined as

$$d_{\tau\varepsilon} = \|\boldsymbol{s}_i^j - \boldsymbol{s}_m^n\|^2 = \begin{cases} |b_j|^2 + |b_n|^2, & if\ i \neq m, j \neq n \\ 2|b_j|^2, & if\ i \neq m, j = n \\ |b_j - b_n|^2, & if\ i = m, j \neq n \\ 0. & if\ i = m, j = n \end{cases} \tag{25}$$

Furthermore, in the scenario of $N_t \gg N_r$, we can have the approximation of

$$\zeta = \sqrt{\frac{\theta N_r}{tr\left[\left(\boldsymbol{H}\boldsymbol{H}^H \right)^{-1} \right]}} \approx \sqrt{\theta N_t}. \tag{26}$$

Then, ζ is independent of $\boldsymbol{H}$ and, hence, $P\left(\boldsymbol{s}_i^j \mapsto \boldsymbol{s}_m^n \,|\boldsymbol{H} \right)$ is independent of the realizations of $\boldsymbol{H}$. Consequently, we have the average or unconditional PEP given by

$$\mathbb{E}_{\boldsymbol{H}} \left\{ P\left(\boldsymbol{s}_i^j \mapsto \boldsymbol{s}_m^n \,|\boldsymbol{H} \right) \right\} \approx Q\left(\sqrt{\frac{\theta N_t d_{\tau\varepsilon}}{2\sigma_u^2}} \right). \tag{27}$$

Finally, the receiver's BER union-bound can be obtained by substituting (27) into (22).

B. Bit Error Rate of the Eavesdropper

The aim of secure communications is to restrict eavesdropper's decoding capability. If available, accurate BER or, at least, a BER lower-bound of the eavesdropper would be very helpful for a transmitter to evaluate the secrecy level and to optimize its transmission parameters. However, in digital communications, the existing methodologies for BER analysis are mostly for BER upper-bounds, very few references have analyzed the BER lower-bounds of digital communication schemes. Nevertheless, the union-bound is a well-known very tight BER upper-bound, especially in medium to high SNR regions. Due to the above-mentioned, in this section, we still introduce the union-bound technique to derive an expression for the eavesdropper's BER, which can be expressed as

$$\begin{aligned} P_e(E) \leq & \frac{1}{kMN_r} \sum_{\boldsymbol{s}_i^j \in \mathcal{S}} \sum_{\boldsymbol{s}_m^n \in \mathcal{S}, \boldsymbol{s}_m^n \neq \boldsymbol{s}_i^j} e(\boldsymbol{s}_i^j, \boldsymbol{s}_m^n) \\ & \times \mathbb{E}_{\boldsymbol{H}, \boldsymbol{H}_e} \left\{ P\left(\boldsymbol{s}_i^j \mapsto \boldsymbol{s}_m^n \,|\boldsymbol{H}, \boldsymbol{H}_e \right) \right\} \end{aligned} \tag{28}$$

where the expectation is taken over both $\boldsymbol{H}$ and $\boldsymbol{H}_e$. Due to the fact that $\boldsymbol{H}$ is independent of $\boldsymbol{H}_e$, we have $\mathbb{E}_{\boldsymbol{H}, \boldsymbol{H}_e}\{P(\boldsymbol{s}_i^j \mapsto \boldsymbol{s}_m^n \,|\boldsymbol{H}, \boldsymbol{H}_e\,)\} = \mathbb{E}_{\boldsymbol{H}} \mathbb{E}_{\boldsymbol{H}_e}\{P(\boldsymbol{s}_i^j \mapsto \boldsymbol{s}_m^n \,|\boldsymbol{H}, \boldsymbol{H}_e\,)\}$. Therefore, we can average $P(\boldsymbol{s}_i^j \mapsto \boldsymbol{s}_m^n \,|\boldsymbol{H}, \boldsymbol{H}_e\,)$ first over $\boldsymbol{H}_e$ and then over $\boldsymbol{H}$ as follows.

According to the ML detector formulated in (20), and assuming that $N_t \gg N_r, N_e$, the conditional PEP in (28) can be derived as

$$\begin{aligned} & P(\boldsymbol{s}_i^j \mapsto \boldsymbol{s}_m^n \,|\boldsymbol{H}, \boldsymbol{H}_e\,) \\ & = P\left\{ \|\, \boldsymbol{z} - \boldsymbol{H}_e \bar{\boldsymbol{P}} \boldsymbol{s}_i^j \,\|^2 > \|\, \boldsymbol{z} - \boldsymbol{H}_e \bar{\boldsymbol{P}} \boldsymbol{s}_m^n \,\|^2 \right\} \\ & = P\left\{ A + B > \frac{1}{2} \|\, \boldsymbol{H}_e \bar{\boldsymbol{P}} (\boldsymbol{s}_i^j - \boldsymbol{s}_m^n) \,\|^2 \right\} \end{aligned} \tag{29}$$

where we define $A = \Re[\boldsymbol{r}_i^H b_j^H \boldsymbol{V}_0^H \boldsymbol{H}_e^H \boldsymbol{H}_e \bar{\boldsymbol{P}} (\boldsymbol{s}_i^j - \boldsymbol{s}_m^n)]$ and $B = \Re[\boldsymbol{v}^H \boldsymbol{H}_e \bar{\boldsymbol{P}} (\boldsymbol{s}_i^j - \boldsymbol{s}_m^n)]$. It can be shown that they are the independent Gaussian random variables with the distributions of $A \sim \mathcal{N}\left(0, \frac{\sigma_r^2 |b_j|^2}{2} \|\boldsymbol{G}^H \boldsymbol{H}_e \boldsymbol{f}\|^2 \right)$ and $B \sim \mathcal{N}\left(0, \frac{\sigma_v^2}{2} \|\boldsymbol{H}_e \boldsymbol{f}\|^2 \right)$, respectively, where $\boldsymbol{G} = \boldsymbol{H}_e \boldsymbol{V}_0$ and $\boldsymbol{f} = \bar{\boldsymbol{P}} (\boldsymbol{s}_i^j - \boldsymbol{s}_m^n)$. Consequently, using the Gaussian Q-function, we can express the conditional PEP of (29) as

$$\begin{aligned} & P(\boldsymbol{s}_i^j \mapsto \boldsymbol{s}_m^n \,|\boldsymbol{H}, \boldsymbol{H}_e\,) \\ & = Q\left(\frac{\|\boldsymbol{H}_e \boldsymbol{f}\|^2}{\sqrt{2[|b_j|^2 \sigma_r^2 \left\| \boldsymbol{G}^H \boldsymbol{H}_e \boldsymbol{f} \right\|^2 + \sigma_v^2 \|\boldsymbol{H}_e \boldsymbol{f}\|^2]}} \right) \\ & \overset{a}{\approx} Q\left(\sqrt{\frac{\|\boldsymbol{H}_e \boldsymbol{f}\|^2}{2\left[|b_j|^2 (1-\theta) + \sigma_v^2 \right]}} \right) \end{aligned} \tag{30}$$

where the relation 'a' holds due to $\lim_{N_t \to \infty} \sigma_r^2 \boldsymbol{G}\boldsymbol{G}^H = \lim_{N_t \to \infty} \frac{1-\theta}{N_t - N_r} \boldsymbol{H}_e \boldsymbol{V}_0 \boldsymbol{V}_0^H \boldsymbol{H}_e^H = (1-\theta)\boldsymbol{I}$.

In (30), let us define $\sigma_{wj}^2 = |b_j|^2 (1-\theta) + \sigma_v^2$, which is in fact the instantaneous value of σ_w^2, when assuming that $\boldsymbol{s}_i^j$ is transmitted. Let us also define $\kappa = \|\boldsymbol{H}_e \boldsymbol{f}\|^2$ with its pdf conditioned on $\boldsymbol{f}$ expressed as $p(\kappa|\boldsymbol{f})$ and the corresponding moment generating function (MGF) as $M_\kappa(t|\boldsymbol{f})$. Then, upon averaging the conditional PEP (30) over $\boldsymbol{H}_e$, we obtain

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{H}_e}\{P(\boldsymbol{s}_i^j \mapsto \boldsymbol{s}_m^n\,|\boldsymbol{H},\boldsymbol{H}_e)\}\\
&\overset{b}{=} \mathbb{E}_{\boldsymbol{H}_e}\left\{\frac{1}{\pi}\int_0^{\frac{\pi}{2}} \exp\left(-\frac{\kappa}{4\sigma_{wj}^2 \sin^2\varphi}\right)d\varphi\right\}\\
&= \frac{1}{\pi}\int_0^{\infty}\int_0^{\frac{\pi}{2}} \exp\left(-\frac{\kappa}{4\sigma_{wj}^2 \sin^2\varphi}\right)p(\kappa|\boldsymbol{f})d\varphi d\kappa\\
&= \frac{1}{\pi}\int_0^{\frac{\pi}{2}} M_\kappa\left(-\frac{1}{4\sigma_{wj}^2 \sin^2\varphi}\,|\boldsymbol{f}\right)d\varphi\\
&\overset{c}{=} \frac{1}{\pi}\int_0^{\frac{\pi}{2}}\left(\frac{4\sigma_{wj}^2 \sin^2\varphi}{4\sigma_{wj}^2 \sin^2\varphi + \|\boldsymbol{f}\|^2}\right)^{N_e} d\varphi. \qquad (31)
\end{aligned}$$

In the above derivations, the equation at 'b' is due to the introduction of the alternative form of the Gaussian Q-function, which is $Q(x) = \pi^{-1}\int_0^{\frac{\pi}{2}} e^{\frac{-x^2}{2\sin^2\varphi}}d\varphi$, while the equation at 'c' is because the MGF $M_\kappa(t|\boldsymbol{f})$ over the Rayleigh fading channels $\boldsymbol{H}_e$ can be expressed as [39]

$$M_\kappa(t|\boldsymbol{f}) = \left(\frac{1}{1-\|\boldsymbol{f}\|^2 t}\right)^{N_e}. \qquad (32)$$

When assuming that $N_t \gg N_r$, we can show that $\lim\limits_{N_t\to\infty} \bar{\boldsymbol{P}}^H\bar{\boldsymbol{P}} = \lim\limits_{N_t\to\infty} \zeta^2\left(\boldsymbol{H}\boldsymbol{H}^H\right)^{-1} = \lim\limits_{N_a\to\infty} \frac{\theta N_b(\boldsymbol{H}\boldsymbol{H}^H)^{-1}}{tr[(\boldsymbol{H}\boldsymbol{H}^H)^{-1}]} = \theta\boldsymbol{I}$. Then, we have

$$\|\boldsymbol{f}\|^2 = (\boldsymbol{s}_i^j - \boldsymbol{s}_m^n)^H\bar{\boldsymbol{P}}^H\bar{\boldsymbol{P}}(\boldsymbol{s}_i^j - \boldsymbol{s}_m^n) \approx \theta d_{\tau\varepsilon}. \qquad (33)$$

Substituting (33) into (31), we obtain the average PEP as

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{H},\boldsymbol{H}_e}\{P(\boldsymbol{s}_i^j \mapsto \boldsymbol{s}_m^n\,|\boldsymbol{H},\boldsymbol{H}_e)\}\\
&\approx \frac{1}{\pi}\int_0^{\frac{\pi}{2}}\left(\frac{4\sigma_{wj}^2 \sin^2\varphi}{4\sigma_{wj}^2 \sin^2\varphi + \theta d_{\tau\varepsilon}}\right)^{N_e} d\varphi. \qquad (34)
\end{aligned}$$

Finally, the approximated BER upper-bound of the eavesdropper can be obtained by substituting (34) into (28).

V. Secrecy Capacity Analysis and Optimal Power-Allocation

After analyzing the BER of the desired receiver and the eavesdropper, let us now analyze the secrecy rate of the SPSM system and consider the optimal power-allocation. The ergodic secrecy rate of a wiretap channel is given by [24]

$$R_s(\theta) = [R(\theta) - R_E(\theta)]^+ \qquad (35)$$

where $R(\theta)$ and $R_E(\theta)$ are the ergodic rates attainable respectively by the receiver and eavesdropper, when given the power-allocation with a factor θ. The secrecy capacity of the SPSM is given by the maximum of $R_s(\theta)$ achieved by an optimum θ, which is expressed as

$$C_s = \max_{0<\theta\leq 1}\{R_s(\theta)\}, \qquad (36)$$

where the optimal θ for (36) may be different for different SNRs.

A. Exact Secrecy Rate for Any N_t

Let us denote $\boldsymbol{s}_{\tau\varepsilon} = \boldsymbol{s}_\tau - \boldsymbol{s}_{\boldsymbol{\varepsilon}}$ and, hence, $\|\boldsymbol{s}_{\tau\varepsilon}\|^2 = d_{\tau\varepsilon}$. By following the analysis in [33], the ergodic rate attained by the receiver can be expressed as

$$\begin{aligned}
R(\theta) &= \log(MN_r) - \frac{1}{MN_r}\sum_{\tau=1}^{MN_r}\\
&\times \mathbb{E}_{\boldsymbol{H},\boldsymbol{u}}\left\{\log\sum_{\varepsilon=1}^{MN_r}\exp\left(\frac{-\|\zeta\boldsymbol{s}_{\tau\varepsilon}+\boldsymbol{u}\|^2+\|\boldsymbol{u}\|^2}{\sigma_u^2}\right)\right\}. \qquad (37)
\end{aligned}$$

From (15) we know that, when N_t is small, $\boldsymbol{w}$ is a colored noise vector with an autocorrelation matric $\boldsymbol{\Sigma}$ given by (16). According to [32], $\boldsymbol{w}$ can be whitened by multiplying $\boldsymbol{\Sigma}^{-1/2}$ on $\boldsymbol{z}$ of (15) without affecting the mutual information, yielding

$$\boldsymbol{z'} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{z} = \boldsymbol{G}\bar{\boldsymbol{P}}\boldsymbol{s}_i^j + \boldsymbol{w'} \qquad (38)$$

where $\boldsymbol{G} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{H}_e$ and $\boldsymbol{w'} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{w}$ becomes a standard white Gaussian noise vector, whose elements are distributed with zero mean and unit variance. In this case, similarly to (37), we can express the ergodic rate attainable by the eavesdropper as

$$\begin{aligned}
R_E(\theta) &= \log(MN_r) - \frac{1}{MN_r}\sum_{\tau=1}^{MN_r}\\
&\times \mathbb{E}_{\boldsymbol{G},\boldsymbol{w'}}\left\{\log\sum_{\varepsilon=1}^{MN_r}\exp\left(-\|\boldsymbol{G}\bar{\boldsymbol{P}}\boldsymbol{s}_{\tau\varepsilon}+\boldsymbol{w'}\|^2+\|\boldsymbol{w'}\|^2\right)\right\}. \qquad (39)
\end{aligned}$$

Finally, the secrecy rate $R_s(\theta)$ achievable by the receiver can be obtained by substituting (37) and (39) into (35).

However, as seen in (37) and (39), neither $R(\theta)$ nor $R_E(\theta)$ are in the closed forms, and it is hard to further simplify them, as the expectations involve some highly complicated integrations related to several random matrices and vectors. Consequently, we are unable to obtain an analytical formula for C_s in (36), which is provided by an optimum θ. Previously in [4], we have evaluated the expectations of $\mathbb{E}_{(\cdot)}$ in $R(\theta)$ and $R_E(\theta)$ via Monte-Carlo simulations for every θ value. Then, the near-optimal θ for $R_s(\theta)$ is obtained via exhaustive search. Our studies show that $R_s(\theta)$ found in this way is close to the secrecy capacity C_s. However, when the input-output dimensions become very large, the Monte Carlo simulation approach will involve a prohibitive computational complexity and become not feasible [29]. Hence, below we derive the approximate but closed-form formulas for $R(\theta)$, $R_E(\theta)$ and $R_s(\theta)$, from which the optimal θ achieving the secrecy capacity C_s can then be obtained with significantly reduced computations, in comparison with the Monte Carlo simulations.

B. Approximate Secrecy Rate for Large N_t

As shown in (26), when $N_t \gg N_r$, ζ is independent of $\boldsymbol{H}$. Hence, the expectation $\mathbb{E}_{(\cdot)}\{\cdot\}$ in (37) is only with respect to $\boldsymbol{u}$. Consequently, we can obtain a lower-bound for $R(\theta)$, which is stated in the following theorem.

Theorem 1: A lower-bound of $R(\theta)$ in the case of $N_t \gg N_r, N_e$ is given by

$$\begin{aligned}R^{(L)}(\theta) =& \log\left(MN_r\right) - N_r\left(\frac{1}{\ln 2}-1\right)\\&-\frac{1}{MN_r}\sum_{\tau=1}^{MN_r}\log\sum_{\varepsilon=1}^{MN_r}\exp\left(-\frac{\theta N_t d_{\tau\varepsilon}}{2\sigma_u^2}\right).\end{aligned} \quad (40)$$

Proof 1: See Appendix A.

For the lower-bound of $R_E(\theta)$, we have known from (19) that $\boldsymbol{w}$ converges to a white Gaussian noise vector with a covariance matrix of $\sigma_w^2\boldsymbol{I}_{N_e}$, where $\sigma_w^2 = (1-\theta)+\sigma_v^2$, when $N_t \gg N_r$. In this case, a lower-bound as well as a upper-bound for the data rate achieved by the eavesdropper can be established, which is given in the following theorem.

Theorem 2: In the case of $N_t \gg N_r, N_e$, a lower-bound of $R_E(\theta)$ is given by

$$\begin{aligned}R_E^{(L)}(\theta) =& \log\left(MN_r\right) - N_e\left(\frac{1}{\ln 2}-1\right)\\&-\frac{1}{MN_r}\sum_{\tau=1}^{MN_r}\log\sum_{\varepsilon=1}^{MN_r}\left(1+\frac{\theta d_{\tau\varepsilon}}{2\sigma_w^2}\right)^{-N_e},\end{aligned} \quad (41)$$

and an upper-bound of $R_E(\theta)$ is given by

$$\begin{aligned}R_E^{(U)}(\theta) =& \log\left(MN_r\right)\\&-\frac{1}{MN_r}\sum_{\tau=1}^{MN_r}\log\sum_{\varepsilon=1}^{MN_r}\exp\left(-\frac{\theta N_e d_{\tau\varepsilon}}{\sigma_w^2}\right).\end{aligned} \quad (42)$$

Proof 2: See Appendix B.

After careful analysis and also with the aid of the numerical demonstrations in [29, 30], we find that $R^{(L)}(\theta)$ and $R_E^{(L)}(\theta)$ can provide very accurate approximation to $R(\theta)$ and $R_E(\theta)$, respectively, after shifting them by a constant. In order to illustrate these, we begin with deriving the limits of $R(\theta)$ in the low SNR and high SNR regions. Since $\gamma = 1/\sigma_u^2$, we can readily derive from (37) that the limits of $R(\theta)$ are given by

$$\lim_{\gamma\to 0} R(\theta) = 0,\ \lim_{\gamma\to+\infty} R(\theta) = \log\left(MN_r\right). \quad (43)$$

Similarly, from (40), we can obtain the limits of $R^{(L)}(\theta)$ as

$$\begin{aligned}&\lim_{\gamma\to 0}\ R^{(L)}(\theta) = -N_r\left(\frac{1}{\ln 2}-1\right),\\&\lim_{\gamma\to+\infty} R^{(L)}(\theta) = \log\left(MN_r\right) - N_r\left(\frac{1}{\ln 2}-1\right).\end{aligned} \quad (44)$$

When comparing the results in (43) with the corresponding ones in (44), we realize that the lower-bound $R^{(L)}(\theta)$ has the same difference from $R(\theta)$ in both the low and high SNR regions. From these results and the fact that $R(\theta)$ is a monotonically increasing function of γ, we can be implied that at any given SNR, the difference between $R(\theta)$ and $R^{(L)}(\theta)$ is approximately a constant of $N_r\left(1/\ln 2 - 1\right)$. Note that this observation has also been reported in [29, 30] for the general MIMO channels. Hence, we are confident that a very good approximation to $R(\theta)$ is $R(\theta) \approx R^{(L)}(\theta) + N_r\left(\frac{1}{\ln 2}-1\right)$. In other words, $R(\theta)$ can be approximately expressed as

$$\hat{R}(\theta) \approx \log\left(MN_r\right) - \frac{1}{MN_r}\sum_{\tau=1}^{MN_r}\log\sum_{\varepsilon=1}^{MN_r}\exp\left(-\frac{\gamma\theta N_t d_{\tau\varepsilon}}{2}\right). \quad (45)$$

Similarly, $R_E(\theta)$ can be approximated by

$$\begin{aligned}\hat{R}_E(\theta) \approx& \log\left(MN_r\right) - \frac{1}{MN_r}\sum_{\tau=1}^{MN_r}\\&\times\log\sum_{\varepsilon=1}^{MN_r}\left(1+\frac{\gamma\theta d_{\tau\varepsilon}}{2[\gamma(1-\theta)+1]}\right)^{-N_e}.\end{aligned} \quad (46)$$

Furthermore, when MPSK is employed as the APM, then, exploiting the symmetric property of the constellations, (45) and (46) can be simplified to

$$\hat{R}_{PSK}(\theta) \approx \log\left(MN_r\right) - \log\sum_{\varepsilon=1}^{MN_r}\exp\left(-\frac{\gamma\theta N_t d_{1\varepsilon}}{2}\right), \quad (47)$$

$$\begin{aligned}\hat{R}_{E-PSK}(\theta) \approx& \log\left(MN_r\right)\\&-\log\sum_{\varepsilon=1}^{MN_r}\left(1+\frac{\gamma\theta d_{1\varepsilon}}{2[\gamma(1-\theta)+1]}\right)^{-N_e}.\end{aligned} \quad (48)$$

Subtracting (46) from (45), we can obtain an *approximation* to $R_s(\theta)$ as

$$\hat{R}_s(\theta) \approx \frac{1}{MN_r}\left(F_2 - F_1\right) \quad (49)$$

where, by definition,

$$F_1 = \sum_{\tau=1}^{MN_r}\log\sum_{\varepsilon=1}^{MN_r}\exp\left(-\frac{\gamma\theta N_t d_{\tau\varepsilon}}{2}\right), \quad (50)$$

$$F_2 = \sum_{\tau=1}^{MN_r}\log\sum_{\varepsilon=1}^{MN_r}\left(1+\frac{\gamma\theta d_{\tau\varepsilon}}{2[\gamma(1-\theta)+1]}\right)^{-N_e} \quad (51)$$

Alternatively, by subtracting (42) from (45), we can obtain an *approximated lower-bound* of $R_s(\theta)$, which is

$$R_s^{(L)}(\theta) \approx \frac{1}{MN_r}\left(F_3 - F_1\right) \quad (52)$$

where $F_3 = \sum\limits_{\tau=1}^{MN_r}\log\sum\limits_{\varepsilon=1}^{MN_r}\exp\left(-\frac{\gamma\theta N_e d_{\tau\varepsilon}}{\gamma(1-\theta)+1}\right)$.

Based on the closed-form formulas of $\hat{R}_s(\theta)$ or $R_s^{(L)}(\theta)$, the optimal values of θ for achieving the highest secrecy rate at different SNR can be readily obtained with the aid of simple one-dimensional search. Furthermore, our numerical results in Section VI-C show that, the value of θ obtained via maximizing $\hat{R}_s(\theta)$ is always closer to the optimal θ, than the one obtained via maximizing $R_s^{(L)}(\theta)$. Therefore, in the following discourses, the asymptotic solutions maximizing $R_s(\theta)$ in low and high SNR regions are only analyzed based on $\hat{R}_s(\theta)$ of (49).

Let us first construct a matrix $\boldsymbol{D}$ from the PSM symbol set, whose entries are $\boldsymbol{D}(\tau,\varepsilon) = d_{\tau\varepsilon} = \|\boldsymbol{s}_\tau - \boldsymbol{s}_\varepsilon\|^2$, for $\tau,\varepsilon = 1,\cdots,MN_r$. Explicitly, the main diagonal elements of $\boldsymbol{D}$ are zero. Let the minimum non-zero element of $\boldsymbol{D}$ be expressed as

d_m. Let L_τ be the number of this minimum elements contained in the τ-th row of $\boldsymbol{D}$, while ℓ be the total number of this minimum elements in $\boldsymbol{D}$, i.e, $\ell = \sum_{\tau=1}^{MN_r} L_\tau$. Specifically, when MPSK is employed, the values of L_τ for all τ are identical, which is denoted as $\bar{L}$. Furthermore, we can readily derive that $\bar{L} = (N_r - 1) \times 2$ for BPSK, $\bar{L} = (N_r - 1) \times 4 + 2$ for QPSK, and $\bar{L} = 2$ for M-PSK with $M > 4$. With the aid of the above definitions, we can express the optimum values of θ in low and high SNR regions as follows.

Theorem 3: In the low SNR region ($\gamma \ll 1$), the value of θ maximizing $\hat{R}_s(\theta)$ is $\theta = 1$. In the high SNR region ($\gamma \gg 1$), the value of θ maximizing $\hat{R}_s(\theta)$ is

$$\theta = \frac{2}{\gamma N_t d_m} \ln\left(\frac{\gamma N_t d_m \ell}{2 N_e N_r M} - \ell\right) \tag{53}$$

for M-QAM, and

$$\theta = \frac{2}{\gamma N_t d_m} \ln\left(\frac{\gamma N_t d_m \bar{L}}{2 N_e} - \bar{L}\right) \tag{54}$$

for M-PSK.

Proof 3: See Appendix C.

VI. Performance Results and Discussion

In this section, we provide a range of results in order to demonstrate that: (i) SPSM inherits the low-complexity detection from PSM, while significantly enhances its security performance; (ii) our analytical results in Section IV for evaluation of BER performance is accurate for sufficiently large N_t; and (iii) the secrecy capacity of SPSM for large N_t can be achieved with the optimal power-allocation obtained from the analysis in Section V. In our performance demonstration, $N_r = 4$, i.e., 4-SSK and 16QAM are used unless they are specially mentioned.

A. Low-Complexity Detection and Security Performance

Fig. 1 shows the BER versus SNR performance of the PSM and the SPSM with $N_t = 10$, $\theta = 0.5$. In this example, we deliberately let $N_e = 6 > N_r = 4$, allowing the eavesdropper to use the optimal ML detector in (17), while the receiver detects using SD or JD. From Fig. 1, first, we observe that the eavesdropper is capable of reliably detecting the PSM symbols and achieving a BER of 10^{-5}, at a cost of about 4 dB more SNR than the desired receiver. Second, as in the PSM, the low-complexity SD for the SPSM is highly efficient, which can achieve nearly the same BER performance as the optimal JD. Third, by dividing half-by-half the transmit power for ZFP and AP, the receiver in the SPSM suffers from about 3 dB performance penalty, in comparison with the PSM. However, the eavesdropper's detection performance is severely degraded, with a BER always above 10^{-2} in the considered SNR region. From this observation, we may further infer that, if the transmitter chooses a smaller value of θ, meaning that more power is used for AP, the eavesdropper's BER can be driven to approach 0.5, while the receiver's BER curve has the same shape as that in Fig. 1, but shifts towards the righthand side.

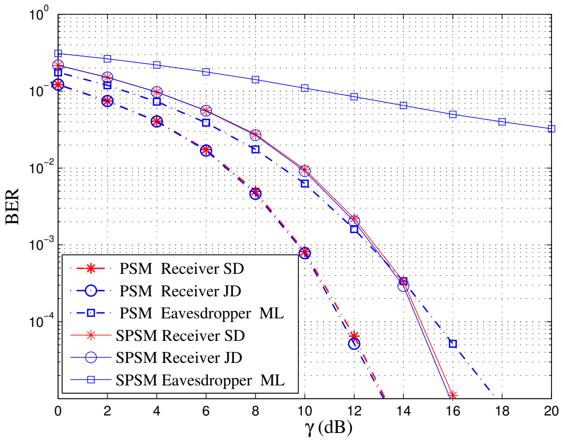


Fig. 1. BER v.s. SNR performance of the PSM and SPSM with $\theta = 0.5$, $N_t = 10$, $N_r = 4$, 16QAM and $N_e = 6$.

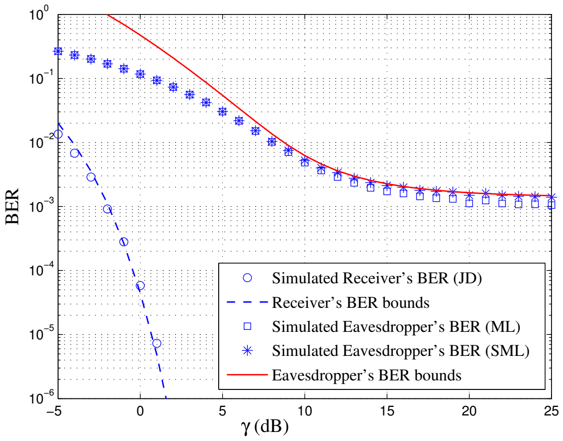


Fig. 2. Comparison between the simulated BER and the analytical BER bounds with $\theta = 0.85$; $N_t = 100$, $N_r = 4$, 16QAM and $N_e = 10$.

B. Validation of Analytical BER Results

Fig. 2 demonstrates the accuracy of our derived BER bounds in Section IV against the Monte Carlo simulations. In this example, we set $N_t = 100$, $N_r = 4$ and $N_e = 10$ in accordance with the assumption of $N_t \gg N_r, N_e$. We assumed that the receiver uses the (joint) ML detection, while the eavesdropper uses either the optimal ML detection of (17) or the simplified ML (SML) detection of (20). Explicitly, Fig. 2 shows that our derived BER bounds are very tight for the receiver over the whole considered SNR region, and for the eavesdropper when the SNR is higher than 5 dB. The eavesdropper's BER conflicts an error floor starting appearing at about 10 dB. According to our discussion below eq. (21), the height of this floor is determined by the value of θ. Therefore, with the aid of the derived BER bounds, a transmitter can adjust the value of θ, in order to make the eavesdropper's detection performance the worst, while satisfying the BER requirement of the intended receiver. In addition to the above, from Fig. 2 we also have the following two observations. First, for the eavesdropper, the SML can achieve a similar BER performance as the optimal

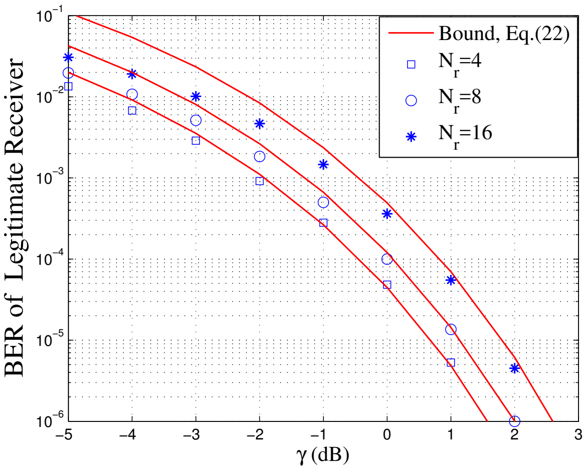


Fig. 3. Comparison between the simulated BER and the analytical BER bound of Eq. (22), when the receiver employs different numbers of antennas, and the other parameters of $\theta = 0.85$, $N_t = 100$, 16QAM and $N_e = 10$.

ML. This observation strongly supports our previous statement in Section IV that $\boldsymbol{w}$ appears white in space in the case of $N_t \gg N_e, N_r$. Second, benefiting from the considerable array gain provided by using a large number of transmit antennas, the receiver is capable of achieving a very low BER in the low SNR region, where the eavesdropper's BER is very high.

In Fig. 3, we investigate the effect of the number of receive antennas on the BER performance of legitimate user, and compare the simulated BER with the upper-bound of Eq. (22) analyzed in Section IV-A. From the results of Fig. 3 we observe that in general the BER bound becomes tight, provided that the SNR is sufficiently high. Specifically, in the considered case of $N_t = 100$, the BER bound is very tight for both $N_r = 4$ and 8, even when the SNR is as low as -2 dB. By contrast, when $N_r = 16$, the bound is tight, provided the SNR is higher than 0 dB. Note that, the reason for the observation that the bound of Eq. (22) becomes looser when increasing N_r is that this bound was derived under the condition of $N_t \gg N_r$, as seen, e.g., in Eq. (26). For a given N_t value, this condition becomes less satisfied, as N_r increases. Consequently, the bound at a given SNR becomes looser, as N_r increases. Note additionally that, as PSM is usually used for downlink transmission (from base-station to mobiles), the number of receive antennas on a mobile unit should not be very large, due to its size constraint.

C. Secrecy Capacity and Power Allocation

Fig. 4 shows the exact secrecy capacity (upper subplot) and its corresponding optimal power-allocation (lower subplot) obtained by Monte Carlo simulations. The results in this figure will also be used to serve as a benchmark in the following figures, where analytical θ values are applied. From this figure, we observe that, even when the eavesdropper has the same number ($N_e = 4$) of antennas as the receiver or has more antennas ($N_e = 10, 20$) than the receiver, C_s is capable of converging to the upper bound of $\log(MN_r)$ bits/s/Hz, as γ increases. This observation explains that, when given sufficiently high transmit power, a proper value of θ can be set by the transmitter, so that the intended receiver can successfully decode the information, while the eavesdropper is unable to due to the artificial interference. As shown in Fig. 4, as N_e increases from 4 to 20, the secrecy capacity of the SPSM decreases and, in order to achieve the secrecy capacity, the transmitter also needs to send a higher fraction of the total transmit power to interfere the eavesdropper.

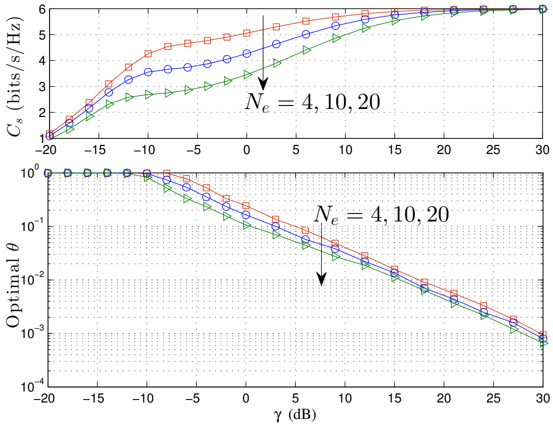


Fig. 4. Simulated secrecy capacity and corresponding optimal θ for the SPSM with the parameters $N_t = 100$, $N_r = 4$, 16QAM and $N_e = 4, 10, 20$.

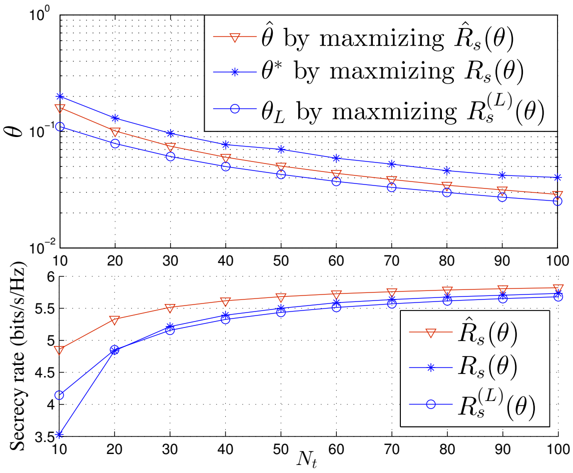


Fig. 5. Comparison between the simulated rate $R_s(\theta)$ and the approximated secrecy rate of $\hat{R}_s(\theta)$ or $R_s^{(L)}(\theta)$, as well as the corresponding power-allocation, when assuming γ=10 dB, $N_r = N_e = 4$, 16QAM, and different number of transmit antennas.

Fig. 5 shows the comparison between the simulated $R_s(\theta)$ and the approximated secrecy rate of $\hat{R}_s(\theta)$ or $R_s^{(L)}(\theta)$, when N_t is increased from $N_t = 10$ to $N_t = 100$. Explicitly, $\hat{R}_s(\theta)$ and $R_s^{(L)}(\theta)$ converge fast to the simulated $R_s(\theta)$, as N_t increases. To be more specific, as shown in the secrecy rate subfigure (lower subfigure), $R_s^{(L)}(\theta)$ converges fast to $R_s(\theta)$ than $\hat{R}_s(\theta)$. By contrast, as shown in the power-allocation subfigure (upper subfigure), the power-allocation obtained from maximizing $\hat{R}_s(\theta)$ is closer to the optimal power-allocation than that obtained from maximizing $R_s^{(L)}(\theta)$. Nevertheless, even though the analytical results in Section V were obtained by assuming $N_t \to \infty$, very accurate results for power-allocation and secrecy rate can be attained, provided that $N_t \geq 20$ for the considered case of $N_r = N_e = 4$. Additionally, as seen in Fig. 5, as N_t increases, more power should be allocated to interfere the eavesdropper, in order to attain an increased secrecy rate. This is because, as N_t increases, the desired receiver can obtain a higher transmit diversity for detection, making it require less transmit power to achieve the same reliability. Hence, the saved power can be used to interfere the eavesdropper.

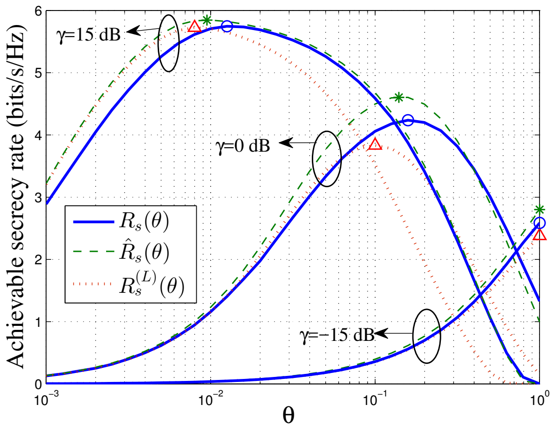


Fig. 6. $R_s(\theta)$, $\hat{R}_s(\theta)$ and $R_s^{(L)}(\theta)$ with respect to different SNRs, when $N_t = 100$, $N_r = 4$, 16QAM and $N_e = 10$.

Fig. 6 shows the functions $\hat{R}_s(\theta)$ and $R_s^{(L)}(\theta)$, as well as the simulated $R_s(\theta)$. Their highest points are respectively marked as $*$, $\triangle$ and $\circ$. First, we can notice that the values of θ achieving the three highest points are close to each other for the three SNR values considered. In particular, all the θ values equal one when the SNR is $\gamma = -15$ dB, which explains that all transmit power should be assigned to the ZFP to guarantee the receiver's detection reliability, when SNR is very low. These results verify that, for the large-scale SPSM systems, the value of θ estimated from maximizing $\hat{R}_s(\theta)$ or $R_s^{(L)}(\theta)$ is a good approximation of the optimal θ.

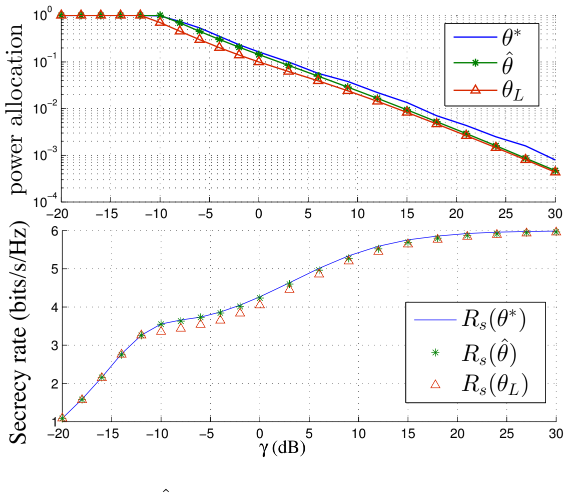


Fig. 7. Values of θ^*, $\hat{\theta}$, and θ_L as well as their corresponding secrecy rate, when $N_t = 100$, $N_r = 4$, 16QAM and $N_e = 10$.

In order to see this more clearly, we plot the optimal θ^* obtained from maximizing $R_s(\theta)$, and the $\hat{\theta}$ and θ_L obtained respectively from maximizing $\hat{R}_s(\theta)$ and $R_s^{(L)}(\theta)$ in the upper subplot of Fig. 7. Their corresponding secrecy rates are shown in the lower subplot of Fig. 7. What we observe is that $\hat{\theta}$ is closer to θ^* than θ_L when $\gamma > -10$ dB, and its corresponding secrecy rate $R_s(\hat{\theta})$ is also very close to $R_s(\theta^*)$, which is actually the secrecy capacity C_s.

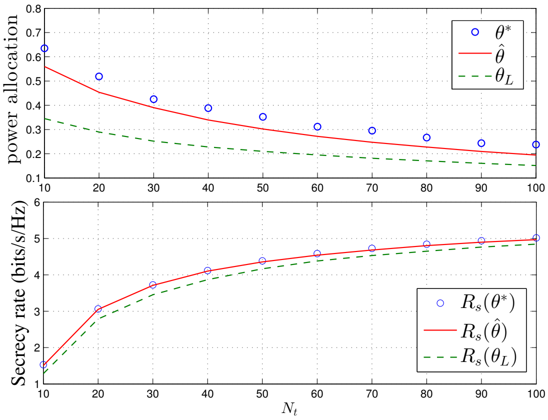


Fig. 8. Values of θ^*, $\hat{\theta}$, and θ_L as well as their corresponding secrecy rate versus N_t changing from 10 to 100, when $N_r = 4$,$N_e = 4$, 16QAM and $\gamma = 0$ dB are assumed.

In order to demonstrate the secrecy rate performance achieved by $\hat{\theta}$ and θ_L, when N_t is moderate, in Fig. 8, we depict the secrecy rate and corresponding power-allocation with respect to N_t that varies from 10 to 100. It can be observed from Fig. 8 that, provided $N_t \geq 30$, the estimated $\hat{\theta}$ becomes close to the optimal θ^*, and the secrecy rate evaluated from Eq.(35) based on $\hat{\theta}$ approaches the secrecy capacity, even when N_t is as small as $N_t = 10$, as shown in the lower subplot of Fig. 8. The above results imply that for medium to large SPSM systems, our proposed near-optimal and yet simple power-allocation strategy is near-optimum for maximizing the secrecy rate, which can be achieved via maximizing $\hat{R}_s(\theta)$.

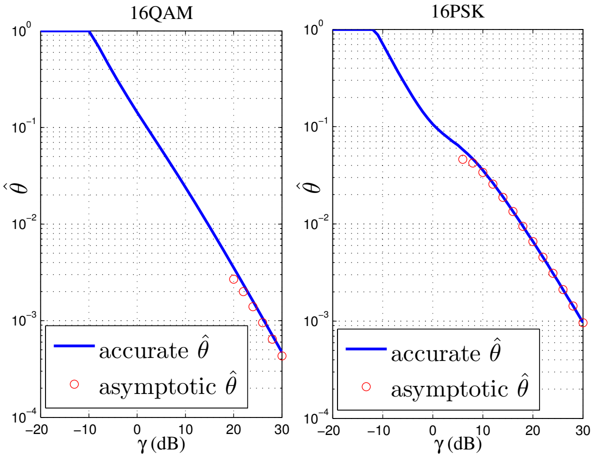


Fig. 9. Comparison between the simulated $\hat{\theta}$ and the high-SNR asymptotic $\hat{\theta}$ for the SPSM systems employing 16QAM (left subplot) or 16PSK (right subplot), when $N_t = 100$, $N_r = 4$ and $N_e = 10$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2841027, IEEE Access

As a matter of fact, in Fig. 7, we have confirmed that $\hat{\theta} = 1$ in low SNR region, which was a statement in Theorem 3 for the low SNR case. In Fig. 9, we further demonstrate the accuracy of the asymptotic $\hat{\theta}$ in high SNR region given in Theorem 3. Explicitly, Fig. 9 shows that, in sufficiently high SNR region, the asymptotic $\hat{\theta}$ agrees well with the accurate results obtained from the exhaustive search. Furthermore, when comparing the two subplots, we can see that the $\hat{\theta}$ estimated from the SPSM employing 16PSK is more accurate than that estimated from the SPSM employing 16QAM. Therefore, when the SNR is sufficiently high, as shown in Fig. 9, the near-optimal power-allocation can be simply obtained from Theorem 3, without requiring any search.

VII. Conclusions

In this paper, we have first analyzed the secrecy capability of the PSM, and explained that it is intrinsically a secrecy or, at least, a LPI transmission scheme when operated in TDD systems. However, when it is operated in the FDD systems, where CSI may have to be fed back from a receiver to its transmitter, making an eavesdropper have opportunity to obtain the CSI, the PSM's security is at risk. In order to enable the PSM to be confidently used in any scenarios, we have therefore proposed the SPSM, which uses a part of power to transmit time-varying interference, in addition to conveying information to the desired receiver. Our studies show that, for the desired receiver, the SPSM employs all the advantages of the PSM, including low-complexity detection. Then, the error performance and secrecy capacity of the SPSM have been comprehensively analyzed under a proposed theoretical framework. A range of expressions have been derived for the upper- and lower-bounds of the BER and secrecy capacity. Furthermore, by both analysis and simulations, we have investigated in detail the optimal power-allocation between information and interference transmission, in order to maximize the secrecy rate. Again, several closed-form formulas have been obtained, which allow us to find the near optimal power-allocation without invoking the Monte-Carlo search of heavy computation, but just relying on simple one-dimension search. Additionally, we have derived the asymptotical expressions for the optimal power-allocation in both low and high SNR regions. All these analytical results are beneficial to the design of SPSM systems, so as to maximize their security.

Appendix A
Proof of Theorem 1

Since ζ in (26) can be treated to be independent of $\boldsymbol{H}$ when $N_t \gg N_r$, the expectation $\mathbb{E}_{\boldsymbol{H},\boldsymbol{u}}$ in (37) only needs to be taken over $\boldsymbol{u}$. Hence, we have

$$\begin{aligned}R(\theta) =&\log\left(MN_r\right) - \frac{N_r}{\ln 2} - \frac{1}{MN_r}\sum_{\tau=1}^{MN_r} \\ &\times \mathbb{E}_{\boldsymbol{u}}\left\{\log\sum_{\varepsilon=1}^{MN_r}\exp\left(\frac{-\left\|\zeta\boldsymbol{s}_{\tau\varepsilon}+\boldsymbol{u}\right\|^2}{\sigma_u^2}\right)\right\} \\ \overset{d}{\geq}&\log\left(MN_r\right) - \frac{N_r}{\ln 2} - \frac{1}{MN_r}\sum_{\tau=1}^{MN_r}\log\sum_{\varepsilon=1}^{MN_r} \\ &\times \mathbb{E}_{\boldsymbol{u}}\left\{\exp\left(\frac{-\left\|\zeta\boldsymbol{s}_{\tau\varepsilon}+\boldsymbol{u}\right\|^2}{\sigma_u^2}\right)\right\}, \qquad (55)\end{aligned}$$

where 'd' holds because of applying the Jensen's inequality to the concave function of $\log(x)$. With the aid of the result of $\mathbb{E}_{\boldsymbol{u}}\left\{\exp\left(\frac{-\|\zeta\boldsymbol{s}_{\tau\varepsilon}+\boldsymbol{u}\|^2}{\sigma_u^2}\right)\right\} = \frac{1}{2^{N_r}}\exp\left(-\frac{\zeta^2\|\boldsymbol{s}_{\tau\varepsilon}\|^2}{2\sigma_u^2}\right)$ [29, Eq. 57], we obtain the lower-bound $R^{(L)}(\theta)$ of (40), which completes the proof.

Appendix B
Proof of Theorem 2

When $N_t \gg N_r, N_e$, $\boldsymbol{w}$ converges to a white Gaussian noise vector, whose elements distribute with zero mean and the variance of σ_w^2. Hence, similar to (37), $R_E(\theta)$ can be expressed as

$$\begin{aligned}R_E(\theta) = \log\left(MN_r\right) - \frac{1}{MN_r}\sum_{\tau=1}^{MN_r} \\ \times \mathbb{E}_{\boldsymbol{H}_e,\boldsymbol{w}}\left\{\log\sum_{\varepsilon=1}^{MN_r}\exp\left(\frac{-\left\|\boldsymbol{H}_e\bar{\boldsymbol{P}}\boldsymbol{s}_{\tau\varepsilon}+\boldsymbol{w}\right\|^2+\left\|\boldsymbol{w}\right\|^2}{\sigma_w^2}\right)\right\}.\end{aligned} \qquad (56)$$

Furthermore, since $\mathbb{E}_{\boldsymbol{H}_e,\boldsymbol{w}}\{\bullet\} = \mathbb{E}_{\boldsymbol{H}_e}\mathbb{E}_{\boldsymbol{w}}\{\bullet\}$, by the same steps for (55), we can arrive at

$$\begin{aligned}R_E(\theta) \geq \log\left(MN_r\right) - N_e\left(\frac{1}{\ln 2}-1\right) \\ - \frac{1}{MN_r}\sum_{\tau=1}^{MN_r}\log\sum_{\varepsilon=1}^{MN_r}\mathbb{E}_{\boldsymbol{H}_e}\left\{\exp\left(-\frac{1}{2\sigma_w^2}\left\|\boldsymbol{H}_e\bar{\boldsymbol{P}}\boldsymbol{s}_{\tau\varepsilon}\right\|^2\right)\right\}.\end{aligned} \qquad (57)$$

Now let us apply the result of [29, eq. 63], which yields

$$\mathbb{E}_{\boldsymbol{H}_e}\left\{\exp\left(-\frac{1}{2\sigma_w^2}\left\|\boldsymbol{H}_e\bar{\boldsymbol{P}}\boldsymbol{s}_{\tau\varepsilon}\right\|^2\right)\right\} = \left(1+\frac{1}{2\sigma_w^2}\left\|\bar{\boldsymbol{P}}\boldsymbol{s}_{\tau\varepsilon}\right\|^2\right)^{-N_e} \qquad (58)$$

where for $N_t \gg N_r$ we have the approximation of

$$\bar{\boldsymbol{P}}^H\bar{\boldsymbol{P}} = \zeta^2\left(\boldsymbol{H}\boldsymbol{H}^H\right)^{-1} \approx \frac{\zeta^2}{N_t}\boldsymbol{I}_{N_r} = \theta\boldsymbol{I}_{N_r}. \qquad (59)$$

Finally, by substituting (59) into (58) and, then, into (57), we obtain the lower-bound of (41).

For deriving the upper-bound, since $\log\left[\sum_{\varepsilon}\exp(x_{\varepsilon})\right]$ is a convex function [41], Jensen's inequality can be applied to (56) to obtain

$$\begin{aligned}R_E(\theta) \leq \log\left(MN_r\right) - \frac{1}{MN_r}\sum_{\tau=1}^{MN_r}\log\sum_{\varepsilon=1}^{MN_r} \\ \times \exp\left(-\mathbb{E}_{\boldsymbol{H}_e}\mathbb{E}_{\boldsymbol{w}}\left\{\frac{\left\|\boldsymbol{H}_e\bar{\boldsymbol{P}}\boldsymbol{s}_{\tau\varepsilon}+\boldsymbol{w}\right\|^2-\|\boldsymbol{w}\|^2}{\sigma_w^2}\right\}\right) \quad (60)\end{aligned}$$

Upon taking the expectation over $\boldsymbol{w}$, we obtain

$$\mathbb{E}_{\boldsymbol{w}}\left\{\frac{\left\|\boldsymbol{H}_e\bar{\boldsymbol{P}}\boldsymbol{s}_{\tau\varepsilon}+\boldsymbol{w}\right\|^2-\|\boldsymbol{w}\|^2}{\sigma_w^2}\right\} = \frac{\left\|\boldsymbol{H}_e\bar{\boldsymbol{P}}\boldsymbol{s}_{\tau\varepsilon}\right\|^2}{\sigma_w^2}. \quad (61)$$

Then, taking the expectation over $\boldsymbol{H}_e$ gives

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{H}_e}\mathbb{E}_{\boldsymbol{w}}\left\{\frac{\left\|\boldsymbol{H}_e\bar{\boldsymbol{P}}\boldsymbol{s}_{\tau\varepsilon}+\boldsymbol{w}\right\|^2-\|\boldsymbol{w}\|^2}{\sigma_w^2}\right\} \\ &= \frac{\boldsymbol{s}_{\tau\varepsilon}^H\bar{\boldsymbol{P}}^H\mathbb{E}_{\boldsymbol{H}_e}\left\{\boldsymbol{H}_e^H\boldsymbol{H}_e\right\}\bar{\boldsymbol{P}}\boldsymbol{s}_{\tau\varepsilon}}{\sigma_w^2} \\ &= \frac{\theta N_e d_{\tau\varepsilon}}{\sigma_w^2}. \quad (62)\end{aligned}$$

Finally, substituting this result into (60), we obtain the upper-bound of $R_E^{(U)}(\theta)$ given in (42).

Appendix C
Proof of Theorem 3

Let us define $\bar{d}_\tau = \frac{1}{MN_r}\sum_{\varepsilon=1}^{MN_r} d_{\tau\varepsilon}$, and $\bar{d} = \frac{1}{MN_r}\sum_{\tau=1}^{MN_r}\bar{d}_\tau = \frac{1}{(MN_r)^2}\sum_{\tau=1}^{MN_r}\sum_{\varepsilon=1}^{MN_r} d_{\tau\varepsilon}$. Then, we can know that $\bar{d}$ is the expectation of $d_{\tau\varepsilon}$, i.e., $\bar{d} = \mathbb{E}\left[\|\boldsymbol{s}_\tau-\boldsymbol{s}_\varepsilon\|^2\right] = \mathbb{E}[\|\boldsymbol{s}_\tau\|^2] + \mathbb{E}[\|\boldsymbol{s}_\varepsilon\|^2] = 2$.

When the SNR is very low, such as $\gamma \ll 1$, we have the approximation $\left(1+\frac{\gamma\theta d_{\tau\varepsilon}}{2[\gamma(1-\theta)+1]}\right)^{-N_e} \approx \left(1+\frac{\gamma\theta d_{\tau\varepsilon}}{2}\right)^{-N_e} \approx 1-\frac{N_e\gamma\theta d_{\tau\varepsilon}}{2}$. Hence F_2 in (51) can be approximated as

$$\begin{aligned}F_2 &\approx \sum_{\tau=1}^{MN_r}\log\sum_{\varepsilon=1}^{MN_r}\left(1-\frac{N_e\gamma\theta d_{\tau\varepsilon}}{2}\right) \\ &= \log\prod_{\tau=1}^{MN_r}\left[MN_r\left(1-\frac{N_e\gamma\theta\bar{d}_\tau}{2}\right)\right] \\ &\approx MN_r\log(MN_r) + \log\left(1-\frac{N_e\gamma\theta}{2}\sum_{\tau=1}^{MN_r}\bar{d}_\tau\right) \\ &= MN_r\log(MN_r) + \log\left(1-\frac{N_e\gamma\theta MN_r\bar{d}}{2}\right) \\ &= MN_r\log(MN_r) + \log\left(1-N_e\gamma\theta MN_r\right). \quad (63)\end{aligned}$$

On the other side, F_1 can be approximated as

$$\begin{aligned}F_1 &\approx \sum_{\tau=1}^{MN_r}\log\sum_{\varepsilon=1}^{MN_r}\left(1-\frac{\gamma\theta N_t d_{\tau\varepsilon}}{2}\right) \\ &= \sum_{\tau=1}^{MN_r}\log\left[MN_r\left(1-\frac{\gamma\theta N_t\bar{d}_\tau}{2}\right)\right] \\ &= MN_r\log(MN_r) + \log\prod_{\tau=1}^{MN_r}\left(1-\frac{N_t\gamma\theta\bar{d}_\tau}{2}\right) \\ &\approx MN_r\log(MN_r) + \log\left(1-\frac{N_t\gamma\theta}{2}\sum_{\tau=1}^{MN_r}\bar{d}_\tau\right) \\ &= MN_r\log(MN_r) + \log\left(1-N_t\gamma\theta MN_r\right). \quad (64)\end{aligned}$$

Substituting (63) and (64) into (49) yields $\hat{R}_s(\theta)$, which is

$$\hat{R}_s(\theta) = \frac{1}{MN_r}\left[\log\left(1-N_e\gamma\theta MN_r\right)-\log\left(1-N_t\gamma\theta MN_r\right)\right]. \quad (65)$$

Taking the first derivative of $\hat{R}_s(\theta)$ with respect to θ, we can easily find that $\frac{\partial\hat{R}_s}{\partial\theta} > 0$, explaining that $\hat{R}_s$ is a monotonically increasing function of θ. Therefore, in the low SNR region, $\hat{R}_s(\theta)$ is maximized by $\theta = 1$.

Considering the high SNR region of $\gamma \gg 1$, the product $\gamma\theta$ may be in one of the three possible cases, which are a) $\gamma\theta \to 0$, b) $\gamma\theta \leq \eta$ with η being a bounded positive number, and c) $\gamma\theta \to \infty$. For the first case, calculating the expression of (49) reveals that $\hat{R}_s(\theta) \to 0$. Hence, any value of θ in this case is definitely not optimal. In the second case, it is implied that $\theta \to 0$ as $\gamma \to \infty$. Finally, for the third case, we obtain $F_1 \to 0$ and $F_2 \approx \sum_{\tau=1}^{MN_r}\log\sum_{\varepsilon=1}^{MN_r}\left(1+\frac{\theta d_{\tau\varepsilon}}{2(1-\theta)}\right)^{-N_e}$. Then according to (49), we have $\hat{R}_s(\theta) = \frac{1}{MN_r}\sum_{\tau=1}^{MN_r}\log\sum_{\varepsilon=1}^{MN_r}\left(1+\frac{\theta d_{\tau\varepsilon}}{2(1-\theta)}\right)^{-N_e}$, which indicates that the value of θ approaching 0 can produce the largest $\hat{R}_s(\theta)$ and be optimum. Therefore, when taking into account of all the three cases, we are confident that the optimal θ in the high SNR region has a very small value, i.e., $\theta \to 0$ as $\gamma \to \infty$. Based on this observation, we have the approximation of $\left(1+\frac{\gamma\theta d_{\tau\varepsilon}}{2[\gamma(1-\theta)+1]}\right)^{-N_e} \approx \left(1+\frac{\theta d_{\tau\varepsilon}}{2}\right)^{-N_e} \approx 1-\frac{N_e\theta d_{\tau\varepsilon}}{2}$. Using this result and following some similar manipulations used in (63), F_2 in the high SNR region can be approximated as

$$F_2 = MN_r\log(MN_r) + \log\left(1-N_e\theta MN_r\right). \quad (66)$$

At the same time, the term of F_1 in (50) can be rewritten as

$$F_1 = \sum_{\tau=1}^{MN_r}\log\left[1+\sum_{\varepsilon=1,\varepsilon\neq\tau}^{MN_r}\exp\left(-\frac{\gamma\theta N_t d_{\tau\varepsilon}}{2}\right)\right]. \quad (67)$$

Let us define $d_{\tau\min} = \min_{\varepsilon\neq\tau,\varepsilon=1,\cdots,MN_r}\{d_{\tau\varepsilon}\}$. Then, we can readily show that, for the rectangular QAM, all $d_{\tau\min}$ for $\tau = 1,\cdots,MN_r$ are identical and equal to d_m. Furthermore, since e^{-x} is a fast decreasing function of positive x, $\sum_{\varepsilon=1,\varepsilon\neq\tau}^{MN_r}\exp\left(-\frac{\gamma\theta N_t d_{\tau\varepsilon}}{2}\right)$ can be approximated by its most significant terms, yielding $\sum\limits_{\varepsilon=1,\varepsilon\neq\tau}^{MN_r} \exp\left(-\frac{\gamma\theta N_t d_{\tau\varepsilon}}{2}\right) \approx L_\tau \exp\left(-\frac{\gamma\theta N_t d_m}{2}\right) \ll 1$. Using these results, F_1 can be approximated as

$$\begin{aligned}
F_1 &\approx \sum_{\tau=1}^{MN_r} \log\left[1 + L_\tau \exp\left(-\frac{\gamma\theta N_t d_m}{2}\right)\right] \\
&= \log \prod_{\tau=1}^{MN_r} \left[1 + L_\tau \exp\left(-\frac{\gamma\theta N_t d_m}{2}\right)\right] \\
&\approx \log\left[1 + \sum_{\tau=1}^{MN_r} L_\tau \exp\left(-\frac{\gamma\theta N_t d_m}{2}\right)\right] \\
&= \log\left[1 + \ell \exp\left(-\frac{\gamma\theta N_t d_m}{2}\right)\right] \qquad (68)
\end{aligned}$$

where ℓ is the total number of d_m appearing in $\boldsymbol{D}$, as defined in Section V-B. Consequently, substituting (66) and (68) into (49), the approximation of $\hat{R}_s(\theta)$ in high SNR region is

$$\begin{aligned}
\hat{R}_s &= \log(MN_r) + \frac{1}{MN_r} \\
&\times \left\{\log\left(1 - N_e\theta MN_r\right) - \log\left[1 + \ell\exp\left(-\frac{\gamma\theta N_t d_m}{2}\right)\right]\right\}.
\end{aligned} \qquad (69)$$

It can be shown that $1 - N_e\theta MN_r$ is log-concave, while $\exp\left(-\frac{\gamma\theta N_t d_m}{2}\right)$ is log-convex [41]. Hence, $\hat{R}_s(\theta)$ in (69) is a concave function of θ. Therefore, the maximum of $\hat{R}_s(\theta)$ is achieved, when θ satisfies

$$\begin{aligned}
\frac{\partial \hat{R}_s}{\partial \theta} &= \frac{1}{MN_r} \\
&\times \left[\frac{-N_e N_r M}{1 - \theta N_e N_r M} + \frac{\frac{\gamma N_t d_m \ell}{2}\exp\left(-\frac{\gamma\theta N_t d_m}{2}\right)}{1 + \ell\exp\left(-\frac{\gamma\theta N_t d_m}{2}\right)}\right] = 0.
\end{aligned} \qquad (70)$$

Finally, with the aid of the approximation $1 - \theta N_e N_r M \approx 1$, solving (70) gives the desired result of (53).

By contrast, when MPSK is employed, we have the simplified expressions of $\hat{R}_{PSK}(\theta)$ in (47) and $\hat{R}_{E-PSK}(\theta)$ in (48). Then, following the similar derivations for (66)-(70), we can obtain (54). For brevity, here we omit the details.

REFERENCES

[1] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications Handbook*. McGraw-Hill New York, 1994.

[2] L.-L. Yang, “Transmitter preprocessing aided spatial modulation for multiple-input multiple-output systems,” in *IEEE 73rd Vehicular Technology Conference (VTC Spring)*, pp. 1–5, May 2011.

[3] F. Wu, R. Zhang, L.-L. Yang, and W. Wang, “Transmitter precoding aided spatial modulation for secrecy communications,” *IEEE Trans. on Veh. Tech.*, vol. 65, pp. 467-471, Jan. 2016.

[4] F. Wu, L.-L. Yang, W. Wang, and Z. Kong, “Secret precoding-aided spatial modulation,” *IEEE Commun. Lett.*, vol. 19, no. 9, pp. 1544–1547, Sep. 2015.

[5] F. Shu, L. Xu , J.Z. Wang, W. Zhu, and X.B. Zhou, ”Artificial-noise-aided secure multicast precoding for directional modulation systems,” *IEEE Trans. Veh. Tech.*, Jan. 2018 (Early Access).

[6] F. Shu, Y. Qin, T. Liu, L. Gui, Y. Zhang, J. Li, and Z. Han, ”Low-complexity and high resolution DOA estimation for hybrid analog and digital massive MIMO receive array”, *IEEE Trans. on Communications*, Feb. 2018 (Early Access).

[7] F. Shu, X.M. Wu, J. Hu, J. Li, R. Chen and J. Wang, ”Secure and precise wireless transmission for random-subcarrier-selection-based directional modulation transmit antenna array,” *IEEE Journal on Selected Areas in Communications*, Mar. 2018 (Early Access).

[8] J. Hu, S. Yan, F. Shu, J. Wang, J. Li, and Y. Zhang, ”Artificial-noise-aided secure transmission with directional modulation based on random frequency diverse arrays,” *IEEE Access*, vol. 5, pp.1658-1667, Feb. 2017.

[9] M. Di Renzo and H. Haas, “Space shift keying (SSK) modulation with partial channel state information: Optimal detector and performance analysis over fading channels,” *IEEE Trans. on Commun.*, vol. 58, no. 11, pp. 3196–3210, Nov. 2010.

[10] R. Zhang, L.-L. Yang, and L. Hanzo, “Generalised pre-coding aided spatial modulation,” *IEEE Trans. on Wireless Commun.*, vol. 12, no. 1, pp. 5434–5443, Nov. 2013.

[11] A. Stavridis, D. Basnayaka, S. Sinanovic, etc., “A virtual MIMO dual-hop architecture based on hybrid spatial modulation,” *IEEE Trans. on Commun.*, vol. 62, no. 9, pp. 3161–3179, Sep. 2014.

[12] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura and L. Hanzo, “Spatial modulation for generalized MIMO: Challenges, opportunities, and implementation,” *Proc. of the IEEE*, vol. 102, no. 1, pp. 56-103, Jan. 2014.

[13] P. Yang, M. Di Renzo, Y. Xiao, S. Li and L. Hanzo, “Design guidelines for spatial modulation,” *IEEE Comm. Surveys & Tutorials*, vol. 17, no. 1, pp. 6-26, Firstquarter 2015.

[14] N. S. Perovic, P. Liu, M. Di Renzo and A. Springer, “Receive spatial modulation for LOS mmWave communications based on TX beamforming,” *IEEE Comm. Letters*, vol. 21, no. 4, pp. 921-924, Apr. 2017.

[15] M. Maleki and K. Mohamed-Pour, “Transmit precoding aided spatial modulation for multi-user correlated large-scale MIMO channels,” in *the 8th International Symposium on Telecomms (IST)*, Tehran, 2016, pp. 337-342.

[16] S. Sinanovic, N. Serafimovski, M. Di Renzo, and H. Haas,“Secrecy capacity of space keying with two antennas,” in *Proc. IEEE VTC-Fall*, Sep. 2012.

[17] S. Sinanovic, M. Di Renzo, and H. Haas, “Secrecy rate of time switched transmit diversity system,” in *Proc. IEEE Veh. Technol. Conf.(VTC Spring)*, Yokohama, Japan, May 2011, pp. 1-5.

[18] X. Guan, Y. Cai, and W. Yang, “On the secrecy mutual information of spatial modulation with finite alphabet,” in *Proc. IEEE WCSP*, Huangshan, China, Oct. 2012, pp. 1-4.

[19] Z. Huang, Z. Gao and L. Sun, “Anti-eavesdropping scheme based on quadrature spatial modulation,” *IEEE Communications Letters*, vol. 21, no. 3, pp. 532-535, March 2017.

[20] L. Wang, S. Bashar, Y. Wei and R. Li, “Secrecy enhancement analysis against unknown eavesdropping in spatial modulation,” *IEEE Communications Letters*, vol. 19, no. 8, pp. 1351-1354, Aug. 2015.

[21] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, etc, “Spatial modulation for generalized MIMO: Challenges, opportunities and implementation,” *Proceedings of the IEEE*, vol. 102, no.1, pp. 56–103, Jan. 2014.

[22] Y. Chen, L. Wang, Z. Zhao, M. Ma and B. Jiao, “Secure multiuser MIMO downlink transmission via precoding-aided spatial modulation,” *IEEE Communications Letters*, vol. 20, no. 6, pp. 1116-1119, June 2016.

[23] S. Goel and R. Negi, “Guaranteeing secrecy using artificial noise,” *IEEE Trans. on Wireless Commun.*, vol. 7, no.6, pp. 2180–2189, Jun. 2008.

[24] X. Zhou and M. McKay, “Secure transmission with artificial noise over fading channels: Achievable rate and optimal power allocation,” *IEEE Trans. on Veh. Technol.*, vol. 59, no. 8, pp. 3831–3842, Oct. 2010.

[25] Y. Zhu, Y. Zhou, S. Patel, X. Chen, L. Pang, and Z. Xue, “Artificial noise generated in MIMO scenario: Optimal power design,” *IEEE Signal Process. Lett.*, vol. 20, no. 10, pp. 964–967, Oct. 2013.

[26] S.-H. Tsai and H. V. Poor,“Power allocation for artificial-noise secure MIMO precoding systems”, *IEEE Trans. Signal Process.*, vol. 62, no. 13, pp. 3479–3493, Jul. 2014.

[27] S. Sugiura and L. Hanzo, “On the joint optimization of dispersion matrices and constellations for near-capacity irregular precoded space-time shift keying,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 380–387, Jan. 2013.

[28] S. X. Ng and L. Hanzo, “On the MIMO channel capacity of multidimensional signal sets,” *IEEE Trans. Veh. Technol.*, vol. 55, pp. 528–536, Mar. 2006.

[29] W. Zeng, C. Xiao, M. Wang, and J. Lu, “Linear precoding for finite-alphabet inputs over MIMO fading channels with statistical CSI,” *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 3134–3148, Jun. 2012.

[30] W. Zeng, C. Xiao, and J. Lu, "A low-complexity design of linear precoding for MIMO channels with finite-alphabet inputs," *IEEE Wireless Commun. Lett.*, vol. 1, no. 1, pp. 38–41, Feb. 2012.

[31] S. Bashar, Z. Ding, and C. Xiao, "On the secrecy rate of multi-antenna wiretap channel under finite-alphabet input," *IEEE Commun. Lett.*, vol. 15, no. 5, pp. 527–529, May 2011.

[32] Y. Wu, C. Xiao and Z. Ding, "Linear precoding for finite-alphabet signaling over MIMOME wiretap channel," *IEEE Trans. Veh. Technol.*, vol. 61, no. 6, pp. 2599-2612, Jul. 2012.

[33] R. Zhang, L.-L. Yang, and L. Hanzo, "Error probability and capacity analysis of generalised precoding aided spatial modulation," *IEEE Trans. on Wireless Commun.*, vol. 14, no. 11, pp. 364–375, Jan. 2015.

[34] A. Stavridis, S. Sinanovic, M. Di Renzo, and H. Haas, "Transmit precoding for receive spatial modulation using imperfect channel knowledge," in *IEEE 75th Veh. Technol. Conf. (VTC Spring)*, May, 2012.

[35] F. Wu, W. Wang, H.-M. Wang and Q. Yin, "A unified mathematical model for spatial scrambling based secure wireless transmission and its wiretap method," in *Proc. IEEE WCSP*, Nanjing, China, Oct. 2011, pp. 1–5.

[36] F. Wu, W. Wang, B. Yao, and Q. Yin, "Effective eavesdropping in the artificial noise aided security scheme," in *IEEE/CIC International Conference on Communications in China*, Xi'an, China, Aug. 2013, pp. 214–218.

[37] J. Zhu, R. Schober, and V. K. Bhargava, "Secure transmission in multi-cell massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4766–4781, Sep. 2014.

[38] J. G. Proakis, *Digital Communications (Fourth Edition)*, New York: McGraw-Hill, 2000.

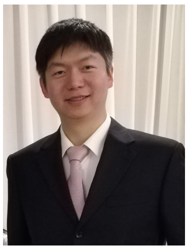
[39] A. Goldsmith, *Wireless communications*, Cambridge, U.K.: Cambridge Univ. Press, 2005.

[40] M. Kiessling, J. Speidel, N. Geng, and M. Reinhardt, "Performance analysis of MIMO maximum likelihood receivers with channel correlation, colored Gaussian noise, and linear prefiltering," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2003, pp. 3026–3030.

[41] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

Feilong Wu is a senior engineer with the Chinese Academy of Space Technology, Xian Branch (CAST, Xian). He received his bachelor degree in electronics and information engineering from Xian University of Science and Technology in July 2009, and his PhD degree in information and communication engineering from Xian Jiaotong University in March 2016. From October 2013 to January 2015, he joined the Next Generation Wireless (NGW) Research Group, University of Southampton, UK, as a visiting PhD student. Since April 2016, he has been with the CAST Xian and working on the design of mobile satellite systems and onboard digital signal processing.

Wenjie Wang (M'10) received his B.S., M.S., and Ph.D. degrees in information and communication engineering from Xi'an Jiaotong University, Xi'an, China, in 1993, 1998, and 2001, respectively. From 2009 to 2010, he was a visiting scholar to the Department of Electrical and Computer Engineering, University of Delaware, Newark, USA. Currently, he is a professor at Xi'an Jiaotong University. His main research interests include information theory, broadband wireless communications, signal processing with application to communication systems, array signal processing and cooperative communications in distributed networks.



Chen Dong received the B.S. degree in electronic information sciences and technology from the University of Science and Technology of China, Hefei, China, in 2004, the MEng degree in pattern recognition and automatic equipment from the University of Chinese Academy of Sciences, Beijing, China, in 2007, and the PhD degree from the University of Southampton, U.K., in 2014. He currently works with the Huawei Device Co., Ltd., China. He was a recipient of a Scholarship under the U.K.-China Scholarships for Excellence Programme and a Best Paper Award at the IEEE VTC 2014. His research interests include applied math, relay system, channel modeling, and cross-layer optimization.

Lie-Liang Yang (M'98, SM'02, F'16) received his BEng degree in communications engineering from Shanghai TieDao University, Shanghai, China in 1988, and his MEng and PhD degrees in communications and electronics from Northern (Beijing) Jiaotong University, Beijing, China in 1991 and 1997, respectively. From June 1997 to December 1997 he was a visiting scientist of the Institute of Radio Engineering and Electronics, Academy of Sciences of the Czech Republic. Since December 1997, he has been with the University of Southampton, United Kingdom, where he is the professor of wireless communications in the School of Electronics and Computer Science. He has research interest in wireless communications, wireless networks and signal processing for wireless communications, as well as molecular communications and nano-networks. He has published 350+ research papers in journals and conference proceedings, authored/co-authored three books and also published several book chapters. The details about his research publications can be found at http://www.mobile.ecs.soton.ac.uk/lly/. He is a fellow of the IET (previously IEE) in the UK, and a distinguished lecturer of the IEEE Vehicular Technology Society. He has served as associate editor to several academic journals, co-organized several special issues, and acted as different roles for conference organization.