Received November 16, 2017, accepted December 25, 2017, date of publication January 8, 2018, date of current version March 9, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2790959

Joint Transmitter–Receiver Spatial Modulation

CHAOWEN LIU[1], LIE-LIANG YANG[2], (Fellow, IEEE), WENJIE WANG[1], (Member, IEEE), AND FASONG WANG[3]

[1]Ministry of Education Key Laboratory for Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an 710049, China
[2]School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K.
[3]School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China

Corresponding author: Lie-Liang Yang (lly@ecs.soton.ac.uk)

The work of C. Liu and W. Wang was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61671366 and in part by the China Scholarship Council under Grant 201506280051. The work of L.-L. Yang was supported by the EPSRC of U.K. under Project EP/P034284/1. The work of F. Wang was supported in part by the NSFC under Grant 61401401, in part by the China Postdoctoral Science Foundation Project under Grant 2015T80779, and in part by the Young Teachers' Special Research Foundation Project of Zhengzhou University under Grant 1521318001.

ABSTRACT We propose and investigate a joint transmitter–receiver spatial modulation (JSM) scheme, which transmits information by jointly exploiting the indices of transmit antenna patterns, the indices of receiver antennas, and amplitude–phase modulation. The proposed JSM is capable of simultaneously achieving transmit diversity, receive diversity, and multiplexing gain. In order to facilitate the implementation of the JSM with different reliability-complexity tradeoff, two types of detectors are introduced, which are the maximum-likelihood detection (MLD) and the proposed sub-optimal detection (SOD), including a two-stage SOD and a three-stage SOD. In order to gain insight into the characteristics of the JSM, we analyze the approximate and asymptotic average bit error probability of the JSM systems employing MLD, when regular-scale and large-scale JSM systems are, respectively, considered. Finally, the performance of the JSM systems is investigated with the aid of both simulation and numerical evaluation of our derived formulas. Our studies show that the JSM employs a high flexibility for implementation and online configuration, so as to adapt to the communications environments and to attain the best possible performance.

INDEX TERMS Multiple-input multiple-output (MIMO), spatial modulation, transmitter preprocessing, signal detection, average bit error probability, asymptotic analysis.

I. INTRODUCTION

By activating one or a fraction of transmit antennas (TAs) in a multiple-input multiple-output (MIMO) system, spatial modulation (SM) [1], [2] employs a range of advantages over the conventional MIMO transmission schemes, such as, VBLAST [3], STBC [4], *etc*. The basic principle of SM is that the degrees of freedom in the spatial domain are directly exploited for information modulation [5]–[11]. For example, space-shift-keying (SSK) modulation [5]–[8] only exploits the indices of TAs for information modulation. It has the advantages of low transceiver implementation-complexity and high energy-efficiency. However, as only the degrees of freedom in the spatial domain are exploited for information modulation, the spectral-efficiency of SSK is relatively low in contrast to the other types of MIMO schemes [9]–[18]. By contrast, SM schemes [9]–[11] make use of both the indices of TAs and the conventional amplitude-phase modulation (APM) for information modulation, in order to achieve a good trade-off between spectral- and energy-efficiency.

Following the principles of SSK and SM, there are various SM-based schemes proposed in recent years [12]–[18]. As some examples, by combining SM with trellis coding and STBC, the trellis-coded SM and STBC-based SM have been respectively proposed to achieve coding gain and transmit diversity [12], [13]. In [14], an adaptive spatial modulation scheme has been designed by proposing a modulation order selection algorithm. Renzo and Haas [15] have analysed in detail the error performance of SM systems, when various fading channels are considered. To improve the spectral-efficiency of SM, the generalised spatial modulation has been proposed in [16], which activates multiple TAs and simultaneously transmits several APM symbols. In [17], a quadrature spatial modulation has been proposed by extending the spatial constellation to both the in-phase and the quadrature dimensions. Additionally, differential SM (DSM) has been proposed in order to free from the channel estimation at receiver [18], which further simplifies the receiver and reduces the pilot overhead for channel estimation.

All the SM schemes mentioned so far exploit the indices of TAs for information modulation. Alternatively, the receive spatial modulation schemes [19] make use of the indices of receive antennas (RAs) for information modulation. As in the receive spatial modulation, transmitter precoding is usually employed, the scheme has originally been referred to as the precoding-aided SM (PSM) [20], which uses precoding, such as zero-forcing (ZF), minimum mean-square error (MMSE), *etc*., to focus transmitted energy on a RA selected by the transmitted information. Similar to the generalised spatial modulation [16], we can design a generalised PSM (GPSM) [21], which simultaneously activates multiple RAs for information delivery. Furthermore, in [22], the capacity and achievable rate of the GPSM systems have been analysed, when either the joint optimal detection or the decoupled sub-optimal detection is employed. In the PSM/GPSM systems [19]–[22], linear precoding is usually employed to facilitate low-complexity implementation of transmitters. By contrast, in [23], a non-linear precoding aided GPSM scheme has been proposed in order to improve the error performance of the GPSM.

More recently, the concept of PSM has been applied and extended by different research works [24]–[28]. Specifically, Zheng [24] have proposed a receive-antenna subset selection algorithm, in order to facilitate the application of PSM in the MIMO systems, where the number of RAs is larger than that of TAs. Following the principles of DSM, the precoding aided DSM (PDSM) has been proposed in [25], so as to reduce the implementation complexity of the PSM detector. Upon applying respectively the PDSM and DSM to the first and second hops of a virtual MIMO [19], a hybrid DSM-aided relay transmission scheme has been proposed and studied in [26]. Furthermore, in order to improve the spectral-efficiency of the GPSM-assisted downlink multiuser MIMO, a generalised space-modulation aided multiuser MIMO transmission scheme has been proposed in [27]. Additionally, in [28], as a variant to the PSM, a virtual spatial modulation scheme has been proposed and analysed, which exploits the indices of the virtual parallel channels, obtained from the singular value decomposition of the MIMO channel, for information modulation.

In the existing PSM or receive spatial modulation schemes, typically, the number of TAs is required to be higher than the number of RAs in order to facilitate the implementation of precoding. Usually, all the TAs are activated to emit preprocessed signals, while all the RAs are utilised to encode information. Hence, conventional PSM schemes do not achieve multiplexing gain as well as receive diversity gain. Additionally, when a big number of TAs is equipped by the transmitter of a PSM-assisted massive MIMO system, the cost for signal preprocessing can become extreme and impractical. In this case, a TA subset based PSM design would be beneficial to the practical implementation, which can also provide multiplexing gain. This is because the indices of the TA subsets can be exploited for transmitting additional information in the space domain. By doing so, the advantages of both the SM and PSM can be attained by a single MIMO system.

Inspired by the above, in this paper, we propose a joint transmitter-receiver spatial modulation (JSM) scheme, which divides TAs into a number of subsets for supporting the preprocessed APM. Therefore, in the JSM, the indices of transmit patterns and the indices of RAs are jointly utilised to modulate information in the space domain, in addition to the APM. To elaborate it further, the main contributions of our paper can be summarised as follows.

1) A novel JSM scheme is designed for signal transmission in the MIMO systems, where the number of TAs is assumed to be significantly higher than that of RAs. In the proposed JSM, the TAs are divided into subsets, and information is jointly transmitted by the indices of transmit patterns, the indices of RAs as well as APM, which are referred to as the TSM, RSM and APM, respectively. In our JSM system, transmit diversity is attained by employing a precoding strategy. By contrast, receive diversity is explicitly achieved by the employment of redundant RAs, with the aid of properly designed detection algorithms. It can be shown that our JSM is capable of achieving transmit diversity gain, receive diversity gain, as well as multiplexing gain.
2) To satisfy different levels of requirements on reliability and complexity, two types of detectors are investigated in this paper. Firstly, we introduce the maximum-likelihood detection (MLD), which is capable of achieving the optimal detection performance but at the highest complexity for implementation. Secondly, we propose two sub-optimal detection (SOD) algorithms, namely the two-stage SOD and three-stage SOD, in order to achieve low-complexity detection. In both the SODs, the RSM symbol is first detected. Following the detection of the RSM symbol, if the two-stage SOD is employed, an optimal search algorithm is implemented to simultaneously detect the TSM and APM symbols at the second stage. By contrast, when the three-stage SOD is employed, the second stage detects the APM symbol, while the TSM symbol is detected at the last stage. Our studies show that the three-stage SOD has the lowest detection complexity, but also the worst detection performance among the three detectors considered.
3) In order to gain the insights into the achievable performance of the JSM systems, we analyse the average bit error probability (ABEP) of the JSM systems with the MLD. In our analysis, both regular-scale MIMO and large-scale MIMO are considered. In the case of regular-scale MIMO, we assume that the number of TAs is limited, and derive the approximate ABEP expressions by deriving the union-bound [29] with the aid of the Gamma approximation (Gamma-Ap). By contrast, the large-scale MIMO assumes an unlimited number of TAs activated. Based on this assumption, we derive the asymptotic ABEP expression by

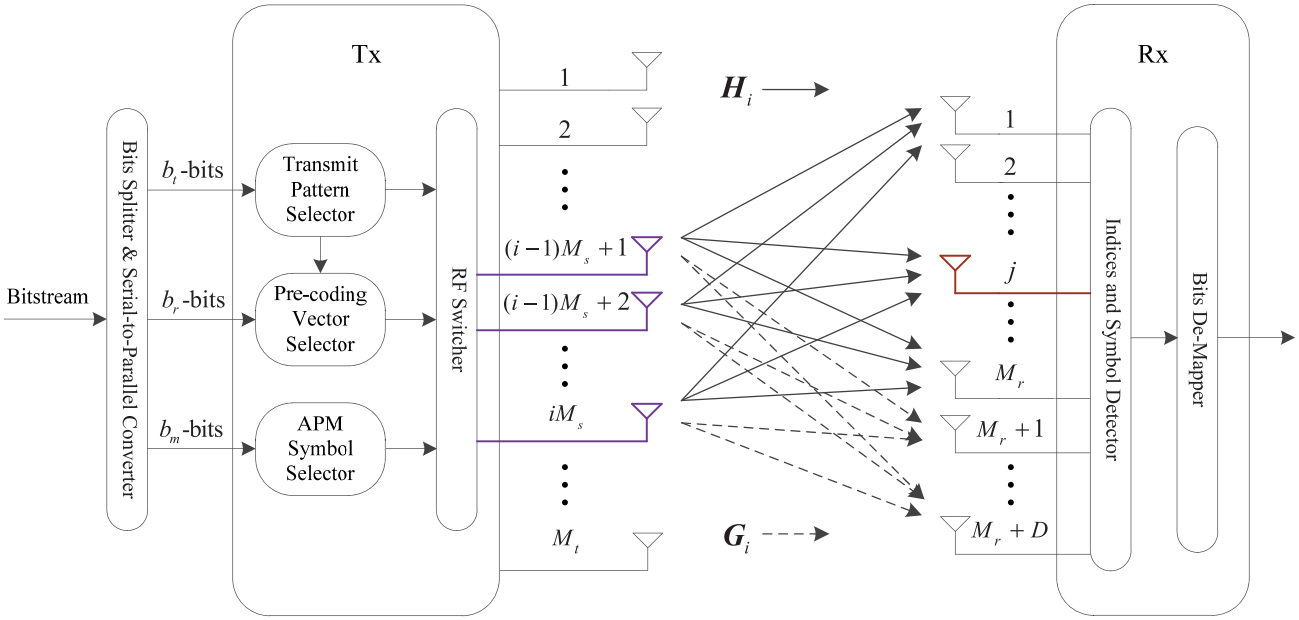


FIGURE 1. An illustration of the proposed JSM system.

introducing both the large-scale approximation and the Gamma-Ap.

The rest of this treatise is organised as follows. In Section II, we describe the implementation procedures of the proposed JSM, and the signalling models. Section III describes the transmitter preprocessing. This is followed by the detection algorithms detailed in Section IV. Section V analyses the ABEP of the JSM systems employing MLD. Simulation and numerical results are depicted and discussed in Section VI, whilst our conclusions are summarised in Section VII.

Notations: Throughout the paper, boldface upper-case and lower-case letters represent matrices and vectors, respectively. The superscripts $\{\}^H$, $\{\}^T$ and $\{\}^*$ stand for the matrix Hermitian transpose, transpose and conjugate, respectively, while $\{\}^{-1}$ is used to denote the inverse of a matrix. $\lfloor x \rfloor$ denotes the maximum integer less than x. $\mathbb{E}[\cdot]$ and $\mathrm{Tr}(\cdot)$ denote respectively the statistical expectation and square-matrix trace operators. $\boldsymbol{I}_M$ represents a identity matrix of M-dimension. $\Re\{\cdot\}$ gets the real part of a complex variable; $\|\cdot\|$ denotes the Euclidean norm of a variable vector. $|\cdot|$ takes the cardinality of a set or the modulus of a complex variable depending on the associated situation.

II. SYSTEM DESCRIPTION AND SIGNALS MODELLING

The JSM system considered consists of a M_t-antenna transmitter and a (M_r+D)-antenna receiver, under the assumption that $M_t/2 \geq M_r \geq 2$, and $D \geq 1$. Let the $(M_r+D)\times M_t$ MIMO channel matrix be expressed as $\boldsymbol{W} = [\boldsymbol{H}^T, \boldsymbol{G}^T]^T$, where $\boldsymbol{H} = [\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_{M_t}] \in \mathbb{C}^{M_r \times M_t}$ and $\boldsymbol{G} = [\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_{M_t}] \in \mathbb{C}^{D\times M_t}$ denote, respectively, the channels from the M_t TAs to the first M_r RAs and that from the M_t TAs to the remaining D RAs. At the receiver, we assume that the first M_r RAs are utilised for information delivery, while the rest D RAs are used to achieve receive diversity, hence referred to as diversity antennas. We assume that, for the purpose of implementation of transmitter preprocessing, the transmitter employs the channel state information (CSI) of $\boldsymbol{H}$, while the whole CSI of $\boldsymbol{W}$ is only known to the receiver. The schematic diagram for the JSM system is illustrated in Fig. 1, which is explained in detail below.

A. PRINCIPLES OF JSM SYSTEMS

In our JSM system, the transmission of an M_m-ary APM symbol S_k ($k \in \{1, 2, \ldots, M_m\}$) is implemented by activating a *subset* of the M_t TAs, which emit the preprocessed signals, in order that the transmitted APM symbol S_k is received only by *one* of the M_r RAs. Therefore, the indices of transmit and receive-antenna patterns are both exploited to bear additional information. To be more specific, the JSM transmission scheme can be elaborated as follows. As shown in Fig. 1, at the transmitter side, the bitstream is first split into data symbols of each containing $b_J = b_t + b_r + b_m$ binary bits. After the serial-to-parallel conversion, each data-symbol is further divided into three parallel sub-symbols of b_t, b_r and b_m bits, which for convenience are referred to as the transmit SM (TSM) symbol, receive SM (RSM) symbol and APM symbol, respectively. Among these three sub-symbols, the TSM symbol of b_t bits determines the transmit pattern given by activating a specified group of TAs. The RSM symbol of b_r bits is used to select the index of one of the M_r RAs. Finally, with the aid of the transmitter preprocessing, the APM symbol of b_m bits is sent by the activated group of TAs, and received by the RA selected from the M_r RAs for RSM as

well as by the other D diversity antennas. Therefore, at the receiver, the objective of detection is to identify the activated transmit-antenna pattern, the activated antenna index among the M_r RSM antennas as well as the APM symbol, in order to detect the b_J $(= b_t + b_r + b_m)$ bits.

In more detail, we assume that the TAs are spatially grouped to form $\lfloor M_t/M_s \rfloor$ transmit patterns, and each transmit pattern contains the same M_s number of antennas. Specifically, the ith, $i \in \{1, 2, \ldots, 2^{b_t}\}$, transmit pattern consists of the antennas with the indices of $\{(i-1)M_s + 1,\ (i-1)M_s + 2, \ldots, iM_s\}$. We assume that $M_r \leq M_s \leq M_t/2$. Let $\mathcal{T}_p$ be a set which includes all the possible transmit patterns. Hence, we have $|\mathcal{T}_p| = \lfloor M_t/M_s \rfloor$, and the number of bits carried by the transmit patterns is given by $b_t = \lfloor \log_2 (M_t/M_s) \rfloor$. Furthermore, given M_r RAs for RSM and constellation size of M_m for APM, the total number of bits transmitted per JSM symbol can be expressed as

$$b_J = \underbrace{\lfloor \log_2 (M_t/M_s) \rfloor}_{b_t} + \underbrace{\lfloor \log_2 (M_r) \rfloor}_{b_r} + \underbrace{\lfloor \log_2 (M_m) \rfloor}_{b_m}. \quad (1)$$

Our JSM can be regarded as a PSM scheme [20] based on subsets of TAs. However, the JSM also uses the transmit patterns to convey extra information, in addition to the information by the RAs' indices and that by the APM in the PSM scheme. Let us use an example to show the possible configurations. Let $M_t = 6$, $M_s = 2$, $M_r = 2$, and $M_m = 4$ (QPSK). Then, we have $|\mathcal{T}_p| = 3$ transmit patterns available, which are $\mathcal{T}_p = \{[1, 2], [3, 4], [5, 6]\}$. Hence, the JSM scheme can deliver in total 4 bits of information per symbol, while the PSM scheme [20] can only convey 3 bits per symbol.

Notice from the above example that, the available transmit patterns can not all be activated for information transmission. Therefore, the transmit patterns can be optimised to improve the performance. In the above example, we can change to use the transmit patterns of $\mathcal{T}_p = \{[1, 2, 3], [4, 5, 6]\}$, which results in transmit diversity as, in this case, $M_s(= 3) > M_r$. Alternatively, we can choose two best from the three transmit patterns, which also yields transmit diversity. In this paper, if the above situation occurs, we simply assume random selection and denote $\mathcal{T}_e = \{\mathcal{T}_p(1), \mathcal{T}_p(2), \mathcal{T}_p(3), \ldots, \mathcal{T}_p(2^{b_t})\}$ as the set of transmit patterns selected. The optimisation of transmit patterns is left to our future research. Note furthermore that, the transmit patterns in our JSM scheme are deliberately selected to be orthogonal. Explicitly, non-orthogonal transmit patterns can also be used. However, using non-orthogonal transmit patterns makes a trade-off for the detection performance due to the added interference.

B. REPRESENTATION OF TRANSMITTED AND RECEIVED JSM SIGNALS

When the ith, $i \in \{1, 2, \ldots, 2^{b_t}\}$, transmit pattern is activated, we express the corresponding channel matrix as $\boldsymbol{W}_i = [\boldsymbol{H}_i^T, \boldsymbol{G}_i^T]^T$, where $\boldsymbol{H}_i = [\boldsymbol{h}_{i,1}, \boldsymbol{h}_{i,2}, \ldots, \boldsymbol{h}_{i,M_s}] \in \mathbb{C}^{M_r \times M_s}$ and $\boldsymbol{G}_i = [\boldsymbol{g}_{i,1}, \boldsymbol{g}_{i,2}, \ldots, \boldsymbol{g}_{i,M_s}] \in \mathbb{C}^{D \times M_s}$, while $\boldsymbol{h}_{i,m_s}$ and $\boldsymbol{g}_{i,m_s}$ $(m_s \in \{1, 2, \ldots, M_s\})$ are the M_r- and D-length channel vectors from the m_sth TA of the ith transmit pattern to the first M_r RAs for RSM, and the remaining D diversity antennas, respectively. Accordingly, the transmitter preprocessing matrix is expressed as $\boldsymbol{P}_i = [\boldsymbol{p}_{i,1}, \boldsymbol{p}_{i,2}, \ldots, \boldsymbol{p}_{i,M_r}]$,[1] where $\boldsymbol{P}_i \in \mathbb{C}^{M_s \times M_r}$ is normalised to satisfy the power constraint of $\mathrm{Tr}(\boldsymbol{P}_i \boldsymbol{P}_i^H) = M_r$. After the preprocessing, the transmitted signal vector can be represented as

$$\boldsymbol{x}_k^{ij} = \boldsymbol{P}_i \boldsymbol{e}_j S_k, \quad (2)$$

where $j \in \{1, 2, \ldots, M_r\}$ denotes the index of the receive antenna selected by the b_r-bits RSM symbol, $\boldsymbol{e}_j$ is the jth column of $\boldsymbol{I}_{M_r}$, and S_k is the APM symbol. We assume that the APM symbols are normalised to have unit average power, i.e., $\mathbb{E}[|S_k|^2] = 1$.

When $\boldsymbol{x}_k^{ij}$ is transmitted over the MIMO channel, as seen in Fig. 1, the observations obtained at the M_r RAs for RSM and the D diversity antennas can be respectively formulated as

$$\boldsymbol{y}_s = \boldsymbol{H}_i \boldsymbol{x}_k^{ij} + \boldsymbol{z}_s = \boldsymbol{H}_i \boldsymbol{P}_i \boldsymbol{e}_j S_k + \boldsymbol{z}_s, \quad (3)$$

$$\boldsymbol{y}_d = \boldsymbol{G}_i \boldsymbol{x}_k^{ij} + \boldsymbol{z}_d = \boldsymbol{G}_i \boldsymbol{P}_i \boldsymbol{e}_j S_k + \boldsymbol{z}_d, \quad (4)$$

where $\boldsymbol{z}_s \in \mathbb{C}^{M_r \times 1}$ and $\boldsymbol{z}_d \in \mathbb{C}^{D \times 1}$ are zero-mean circularly symmetric complex Gaussian noise vectors, whose covariance matrices are $\sigma^2 \boldsymbol{I}_{M_r}$ and $\sigma^2 \boldsymbol{I}_D$, respectively. When combining (3) and (4), we have the presentations of

$$\boldsymbol{y} = \boldsymbol{W}_i \boldsymbol{P}_i \boldsymbol{e}_j S_k + \boldsymbol{z}, \quad (5)$$

where by definition, $\boldsymbol{y} = [\boldsymbol{y}_s^T, \boldsymbol{y}_d^T]^T$, $\boldsymbol{W}_i = [\boldsymbol{H}_i^T, \boldsymbol{G}_i^T]^T$, and $\boldsymbol{z} = [\boldsymbol{z}_s^T, \boldsymbol{z}_d^T]^T$.

Let us below consider the transmitter preprocessing schemes for deriving $\boldsymbol{P}_i$.

III. TRANSMITTER PREPROCESSING FOR JSM SYSTEMS

For a given transmit pattern, such as, the ith transmit pattern, the objective of preprocessing is the same as that for the original PSM system [20]. It is designed to ensure that the output power of the intended RA for RSM is maximised, while the power leakage on the other $(M_r - 1)$ undesired RAs for RSM is minimised. To this end, we directly adopt two canonical transmitter preprocessing strategies considered in [20], namely the transmitter zero-forcing (TZF) preprocessing and transmitter minimum-mean-square-error (TMMSE) preprocessing, while focusing our attention on the novel issues of the JSM systems.

As detailed in [20], when the TZF is employed, the preprocessing matrix is formulated as

$$\boldsymbol{P}_i = \eta_i \boldsymbol{H}_i^H \left(\boldsymbol{H}_i \boldsymbol{H}_i^H \right)^{-1}, \quad (6)$$

[1] Since the transmit pattern varies with the b_t bits of information transmitted, the matrix $\boldsymbol{P}_i$ designed based on the channel state information (CSI) of $\boldsymbol{H}_i$ is different, when different b_t bits of information is transmitted.

where $\eta_i = \sqrt{M_r/\mathrm{Tr}[(\boldsymbol{H}_i\boldsymbol{H}_i^H)^{-1}]}$ is for power normalisation. By contrast, when the TMMSE is employed, we have

$$\boldsymbol{P}_i = \eta_i \boldsymbol{H}_i^H \left(\boldsymbol{H}_i\boldsymbol{H}_i^H + \beta \boldsymbol{I}_{M_r}\right)^{-1}, \tag{7}$$

associated with $\eta_i = \sqrt{M_r/\mathrm{Tr}[(\boldsymbol{H}_i\boldsymbol{H}_i^H + \beta\boldsymbol{I}_{M_r})^{-2}\boldsymbol{H}_i\boldsymbol{H}_i^H]}$. In (7), the regularisation factor β is given by $\beta = M_r\sigma^2$, when TMMSE is employed.

Note that, in general, β in (7) can be set to any positive value, forming a generalised linear transmitter preprocessing (GLTP). In this case, both the TZF and TMMSE can be treated as the special cases of GLTP. From (6) and (7) we are implied that the performance of the TZF and TMMSE converges, as the signal-to-noise ratio (SNR) increases, i.e., as $\sigma^2 \to 0$. Furthermore, as shown in [20] and also the principles of TZF and TMMSE, the BER performance of the TZF is always worse than that of the TMMSE in the low-to-moderate SNR ranges. In other words, the BER given by the TZF constitutes an upper-bound for the BER of the GLTP of (7). Therefore, for brevity, we consider only the TZF in our forthcoming derivation and analysis.

IV. DETECTION ALGORITHMS FOR JSM SYSTEMS

In this section, we propose two types of detectors for the JSM system, which are the optimal maximum-likelihood detection (MLD) and the sub-optimal detection (SOD), when assuming that full CSI is known to the receiver. The difference between the two types of detectors is that the MLD carries out joint detection, while the SOD detects the spatial and APM symbols separately. Hence, the MLD is capable of achieving better error performance than the SOD, but at the cost of higher detection complexity.

A. MAXIMUM-LIKELIHOOD DETECTION

In order to derive the MLD, we first re-formulate the signals observed at the $M_r + D$ RAs. Specifically, when the TZF is employed by the JSM system, by substituting (6) into (3) and (4), the observations from the M_r RAs for RSM and D diversity antennas can be written as

$$\boldsymbol{y}_s = \eta_i \boldsymbol{e}_j S_k + \boldsymbol{z}_s, \tag{8}$$

$$\boldsymbol{y}_d = \boldsymbol{G}_i \boldsymbol{x}_k^{ij} + \boldsymbol{z}_d = \boldsymbol{G}_i \boldsymbol{p}_{i,j} S_k + \boldsymbol{z}_d, \tag{9}$$

where $\boldsymbol{p}_{i,j}$ is the jth column of the preprocessing matrix $\boldsymbol{P}_i$ given by (6).

Given that full CSI is known to the receiver, the optimal MLD provides the estimates to the TSM, RSM and APM symbols by solving the optimisation problem:

$$\begin{aligned}\left\langle \hat{i}, \hat{j}, \hat{k} \right\rangle &= \underset{\tilde{i}\in\mathcal{I},\, \tilde{j}\in\mathcal{J},\, \tilde{k}\in\mathcal{K}}{\arg\min} \left\{ \left\| \boldsymbol{y} - \boldsymbol{W}_{\tilde{i}} \boldsymbol{P}_{\tilde{i}} \boldsymbol{e}_{\tilde{j}} S_{\tilde{k}} \right\|^2 \right\} \\ &= \underset{\tilde{i}\in\mathcal{I},\, \tilde{j}\in\mathcal{J},\, \tilde{k}\in\mathcal{K}}{\arg\max} \left\{ 2\Re\left\{ \boldsymbol{y}^H \boldsymbol{W}_{\tilde{i}} \boldsymbol{p}_{\tilde{i},\tilde{j}} S_{\tilde{k}} \right\} - \left\| \boldsymbol{W}_{\tilde{i}} \boldsymbol{p}_{\tilde{i},\tilde{j}} S_{\tilde{k}} \right\|^2 \right\}, \end{aligned} \tag{10}$$

where $\boldsymbol{y}$ is given by (5) with the components given by (8) and (9), $\mathcal{I} = \{1, 2, \ldots, 2^{b_t}\}$, $\mathcal{J} = \{1, 2, \ldots, 2^{b_r}\}$, and $\mathcal{K} = \{1, 2, \ldots, 2^{b_m}\}$, while $\langle \hat{i}, \hat{j}, \hat{k} \rangle$ represent the estimates to the indices of the TSM, RSM and APM symbols. Upon applying (8) and (9), we can simplify (10) to

$$\begin{aligned}\left\langle \hat{i}, \hat{j}, \hat{k} \right\rangle = \underset{\tilde{i}\in\mathcal{I},\, \tilde{j}\in\mathcal{J},\, \tilde{k}\in\mathcal{K}}{\arg\max} \Bigg\{ &2\Re\left\{ \eta_{\tilde{i}} y_{\tilde{j}}^* S_{\tilde{k}} + \boldsymbol{y}_d^H \boldsymbol{G}_{\tilde{i}} \boldsymbol{p}_{\tilde{i},\tilde{j}} S_{\tilde{k}} \right\} \\ &- \left|\eta_{\tilde{i}} S_{\tilde{k}}\right|^2 - \left\| \boldsymbol{G}_{\tilde{i}} \boldsymbol{p}_{\tilde{i},\tilde{j}} S_{\tilde{k}} \right\|^2 \Bigg\}. \end{aligned} \tag{11}$$

It can be shown that in the general case, the complexity of the MLD based on (10) is $\mathcal{O}(2^{b_t+b_r+b_m} M_s(M_r + D))$. When the TZF is employed, the complexity of the simplified MLD based on (11) is $\mathcal{O}(2^{b_t+b_r+b_m}(M_s D + 1))$.

B. SUB-OPTIMAL DETECTION

From the above analysis, we can see that the complexity of the MLD can be extreme, when $b_J = b_t + b_r + b_m$ is relatively high. Therefore, this section motivates to propose the SOD algorithms of relatively low complexity. Specifically, we propose two SOD algorithms, namely the two-stage SOD and three-stage SOD. As detailed below, both the SODs estimate the RSM symbol at the first stage. Then, based on the estimated RSM symbol, the two-stage SOD jointly estimates the TSM and APM symbols. In order to further reduce the detection complexity, the three-stage SOD also detects the TSM and APM symbols separately, which detects the APM symbol at the second stage, and the TSM symbol at the final stage. Let us now provide the details.

As above-mentioned, for both the SODs, the first stage tries to detect the RSM symbol. This is achieved by identifying the RA that outputs the highest power among the M_r RAs for RSM. Correspondingly, the detection problem can be described as

$$\hat{j} = \underset{\tilde{j}\in\mathcal{J}}{\arg\max} \left\{ \left| y_{\tilde{j}} \right|^2 \right\}. \tag{12}$$

Having detected the RSM symbol, then, for the two-stage SOD, the TSM and APM symbols can be detected by solving the optimisation problem stated as

$$\begin{aligned}\left\langle \hat{i}, \hat{k} \right\rangle = \underset{\tilde{i}\in\mathcal{I},\, \tilde{k}\in\mathcal{K}}{\arg\max} \Bigg\{ &2\Re\left\{ \eta_{\tilde{i}} y_{\hat{j}}^* S_{\tilde{k}} + \boldsymbol{y}_d^H \boldsymbol{G}_{\tilde{i}} \boldsymbol{p}_{\tilde{i},\hat{j}} S_{\tilde{k}} \right\} \\ &- \left|\eta_{\tilde{i}} S_{\tilde{k}}\right|^2 - \left\| \boldsymbol{G}_{\tilde{i}} \boldsymbol{p}_{\tilde{i},\hat{j}} S_{\tilde{k}} \right\|^2 \Bigg\}. \end{aligned} \tag{13}$$

By contrast, in the three-stage SOD, the TSM and APM symbols are separately detected, with the APM symbol detected at the second stage followed by detecting the TSM symbol at the third stage. Specifically, after knowing $\hat{j}$ from (12), we can make use of the observation attained from the $\hat{j}$th RA to detect the APM symbol S_k as

$$\hat{k} = \underset{\tilde{k}\in\mathcal{K}}{\arg\min} \left\{ \left| y_{\hat{j}} - \bar{\eta} S_{\tilde{k}} \right|^2 \right\}, \tag{14}$$

where $\bar{\eta}$ represents the time-averaged value of η_i. With the aid of the related results in [20], we can obtain that $\bar{\eta} = \sqrt{M_s}$. To this point, the receiver has the estimates of $\hat{j}$ and $\hat{k}$, as shown in (12) and (14). Hence, at the third stage, the index of the activated transmit pattern for the TSM symbol can be detected as

$$\begin{aligned}\hat{i} &= \arg\min_{\tilde{i}\in\mathcal{I}} \left\{ \left\| \boldsymbol{y}_d - \boldsymbol{G}_{\tilde{i}} \boldsymbol{p}_{\tilde{i},\hat{j}} S_{\hat{k}} \right\|^2 \right\} \\ &= \arg\max_{\tilde{i}\in\mathcal{I}} \left\{ 2\Re\left\{ \boldsymbol{y}_d^H \boldsymbol{G}_{\tilde{i}} \boldsymbol{p}_{\tilde{i},\hat{j}} S_{\hat{k}} \right\} - \left\| \boldsymbol{G}_{\tilde{i}} \boldsymbol{p}_{\tilde{i},\hat{j}} S_{\hat{k}} \right\|^2 \right\}. \quad (15)\end{aligned}$$

From (12) and (13), we can derive that the computational complexity of the two-stage SOD is $\mathcal{O}(M_r + 2^{b_t+b_m}(M_s D + 1))$, where the first term is contributed by the detection of the RSM symbol, while the second term is by the joint detection of the TSM and APM symbols. Explicitly, the two-stage SOD has lower complexity than the joint MLD of (11), especially in the case when the value of M_r is relatively high. The complexity of the three-stage SOD can be readily analysed from (12), (14), and (15), and is given by $\mathcal{O}(M_r + M_m + 2^{b_t} M_s D)$, which is further reduced in comparison to the two-stage SOD. As we know, the decrease of detection complexity is in general accompanied with the distinct degradation of detection performance, i.e., the increase of BER. However, as demonstrated by our performance results in Section VI, the performance difference between the MLD and two-stage SOD, and that between the two-stage and three-stage SODs are very small in most cases interested.

V. ERROR PERFORMANCE ANALYSIS OF JSM SYSTEMS

In this section, we analyse the average bit error probability (ABEP) of the proposed JSM systems employing the MLD, which gives the upper-bound performance of the JSM systems with other detection algorithms. In our analysis, we assume that the channels from any TAs to any RAs are independent and experience flat Rayleigh fading. Furthermore, the involved channels are independent with respect to the JSM symbol rate, while time invariant within each JSM symbol period. In order to provide comprehensive and insightful analysis, we consider two specific scenarios. In the first scenario, we assume that the number of activated TAs is finite, and derive the approximate ABEP. By contrast, in the second scenario, we derive the asymptotic ABEP of the JSM system, when assuming that the number of activated TAs is infinite.

Our analysis starts with the general formula for the ABEP of the JSM system. With referring to the analysis in [30], the ABEP P_B of the proposed JSM system can be formulated as

$$P_B = \kappa \mathbb{E}_{\boldsymbol{W}} \left[\sum_{i=1}^{2^{b_t}} \sum_{j=1}^{2^{b_r}} \sum_{k=1}^{2^{b_m}} \sum_{\tilde{i}=1}^{2^{b_t}} \sum_{\tilde{j}=1}^{2^{b_r}} \sum_{\tilde{k}=1}^{2^{b_m}} D_{\mathrm{H}}\left(\langle i,j,k\rangle, \langle \tilde{i},\tilde{j},\tilde{k}\rangle \right) \times P_{\mathrm{PE}}\left\{ \langle i,j,k\rangle, \langle \tilde{i},\tilde{j},\tilde{k}\rangle \right\} \right], \quad (16)$$

where $\kappa = (b_J 2^{b_J})^{-1}$, $\mathbb{E}_{\boldsymbol{W}}[\cdot]$ is the expectation with respect to the fading channels $\boldsymbol{W}$, $D_{\mathrm{H}}(\phi, \psi)$ is the Hamming distance between the equivalent bit representations of ϕ and ψ, while $P_{\mathrm{PE}}\{\langle i,j,k\rangle, \langle \tilde{i},\tilde{j},\tilde{k}\rangle\}$ denotes the pairwise error probability (PEP) between the symbols determined by $\langle i,j,k\rangle$ and $\langle \tilde{i},\tilde{j},\tilde{k}\rangle$, which is the probability of detecting $\langle \tilde{i},\tilde{j},\tilde{k}\rangle$, when given that $\langle i,j,k\rangle$ are actually activated. From (16), we can know that simplification of the PEP is very hard and in general, the closed-form formulas for P_B is not available.

In order to obtain simplified but still insightful analytical expressions, below we consider the union-bound [29]. In this case, the ABEP of (16) can be upper-bounded as

$$P_B \le \kappa \sum_{\langle i,j,k\rangle} \sum_{\langle \tilde{i},\tilde{j},\tilde{k}\rangle} D_{\mathrm{H}}\left(\langle i,j,k\rangle, \langle \tilde{i},\tilde{j},\tilde{k}\rangle \right) \times \mathbb{E}_{\boldsymbol{W}}\left[P_{\mathrm{PE}}\left\{ \langle i,j,k\rangle, \langle \tilde{i},\tilde{j},\tilde{k}\rangle \right\} \right], \quad (17)$$

where $\mathbb{E}_{\boldsymbol{W}}[P_{\mathrm{PE}}\{\langle i,j,k\rangle, \langle \tilde{i},\tilde{j},\tilde{k}\rangle\}]$ denotes the average pairwise error probability (APEP), obtained via averaging the PEP over the possible channel realisations of $\boldsymbol{W}$.

For a given $\boldsymbol{W}$, when the MLD of (10) is utilised, the PEP between the symbols determined by $\langle i,j,k\rangle$ and $\langle \tilde{i},\tilde{j},\tilde{k}\rangle$ can be derived as

$$\begin{aligned} &P_{\mathrm{PE}}\left\{ \langle i,j,k\rangle, \langle \tilde{i},\tilde{j},\tilde{k}\rangle \right\} \\ &= P_r\left\{ \left\| \boldsymbol{y} - \boldsymbol{W}_i \boldsymbol{P}_i \boldsymbol{e}_j S_k \right\|^2 > \left\| \boldsymbol{y} - \boldsymbol{W}_{\tilde{i}} \boldsymbol{P}_{\tilde{i}} \boldsymbol{e}_{\tilde{j}} S_{\tilde{k}} \right\|^2 \right\} \\ &= P_r\left\{ 2\Re\left\{ \left(\boldsymbol{W}_{\tilde{i}} \boldsymbol{p}_{\tilde{i},\tilde{j}} S_{\tilde{k}} - \boldsymbol{W}_i \boldsymbol{p}_{i,j} S_k \right)^H \boldsymbol{z} \right\} \right. \\ &\qquad \left. > \left\| \boldsymbol{W}_i \boldsymbol{p}_{i,j} S_k - \boldsymbol{W}_{\tilde{i}} \boldsymbol{p}_{\tilde{i},\tilde{j}} S_{\tilde{k}} \right\|^2 \right\} \\ &= Q\left(\sqrt{ \left\| \boldsymbol{W}_i \boldsymbol{p}_{i,j} S_k - \boldsymbol{W}_{\tilde{i}} \boldsymbol{p}_{\tilde{i},\tilde{j}} S_{\tilde{k}} \right\|^2 / \left(2\sigma^2\right) } \right) \\ &= Q\left(\sqrt{ \frac{\zeta + \xi}{2\sigma^2} } \right), \end{aligned} \quad (18)$$

where $Q(x)$ denotes the Gaussian Q-function defined as $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{t^2}{2}} dt$, $\zeta = \|\eta_i \boldsymbol{e}_j S_k - \eta_{\tilde{i}} \boldsymbol{e}_{\tilde{j}} S_{\tilde{k}}\|^2$ and $\xi = \|\boldsymbol{G}_i \boldsymbol{p}_{i,j} S_k - \boldsymbol{G}_{\tilde{i}} \boldsymbol{p}_{\tilde{i},\tilde{j}} S_{\tilde{k}}\|^2$. When substituting (18) into (17), we have the general APEP expressed as

$$\mathbb{E}_{\boldsymbol{W}}[P_{\mathrm{PE}}\{\langle i,j,k\rangle, \langle \tilde{i},\tilde{j},\tilde{k}\rangle\}] = \mathbb{E}_{\boldsymbol{H},\boldsymbol{G}}\left[Q\left(\sqrt{ \frac{\zeta + \xi}{2\sigma^2} } \right) \right]. \quad (19)$$

In order to simplify (19), we exploit the tight upper-bound introduced in [31] for the Q-function, which is $Q(x) \le \exp(-2x^2)/6 + \exp(-x^2)/12 + \exp(-x^2/2)/4$. Then, the PEP of (18) can be re-formulated as

$$P_{\mathrm{PE}}\left\{ \langle i,j,k\rangle, \langle \tilde{i},\tilde{j},\tilde{k}\rangle \right\} \le \sum_{n=1}^{3} \alpha_n \exp\left(-\frac{\zeta+\xi}{\gamma_n} \right), \quad (20)$$

where, according to the approximation, $\alpha_1 = \frac{1}{6}$, $\alpha_2 = \frac{1}{12}$, and $\alpha_3 = \frac{1}{4}$, while $\gamma_n = \tau_n \sigma^2$, associated with $\tau_1 = 1$, $\tau_2 = 2$,

and $\tau_3 = 4$. Substituting (20) into (19), we have

$$\mathbb{E}_{\boldsymbol{W}}[P_{\mathrm{PE}}\{\langle i,j,k\rangle, \langle \tilde{i},\tilde{j},\tilde{k}\rangle\}] \leq \mathbb{E}_{\boldsymbol{H},\boldsymbol{G}}[\Delta(\zeta,\xi)], \quad (21)$$

where $\Delta(\zeta,\xi)$ is defined as $\Delta(\zeta,\xi) = \sum_{n=1}^{3} \alpha_n e^{-(\zeta+\xi)/\gamma_n}$. Finally, upon substituting (21) into (17), the ABEP of the JSM system satisfies

$$P_B \leq \kappa \sum_{\langle i,j,k\rangle} \sum_{\langle \tilde{i},\tilde{j},\tilde{k}\rangle} D_{\mathrm{H}}\left(\langle i,j,k\rangle, \langle \tilde{i},\tilde{j},\tilde{k}\rangle\right) \mathbb{E}_{\boldsymbol{H},\boldsymbol{G}}[\Delta(\zeta,\xi)]. \quad (22)$$

Let us express $\boldsymbol{G}_i$ and $\boldsymbol{G}_{\tilde{i}}$ respectively as $\boldsymbol{G}_i = \boldsymbol{G}\boldsymbol{A}_i$ and $\boldsymbol{G}_{\tilde{i}} = \boldsymbol{G}\boldsymbol{A}_{\tilde{i}}$, where $\boldsymbol{A}_i$ and $\boldsymbol{A}_{\tilde{i}}$ are constituted by the columns chosen from the identity matrix $\boldsymbol{I}_{M_t}$, based on the ith and $\tilde{i}$th transmit-antenna patterns. Using the assumption that $\boldsymbol{H}$ and $\boldsymbol{G}$ are independent, and with the aid of the moment generating function (MGF) approach [29], the expectation at the right-hand-side (RHS) of the inequality (21) can be simplified as

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{H},\boldsymbol{G}}[\Delta(\zeta,\xi)] \\
&= \mathbb{E}_{\boldsymbol{H}}\left[\sum_{n=1}^{3} \alpha_n \exp\left(-\frac{\zeta}{\gamma_n}\right) \mathbb{E}_{\boldsymbol{G}}\left[\exp\left(-\frac{\xi}{\gamma_n}\right)\right]\right] \\
&= \mathbb{E}_{\boldsymbol{H}}\left[\sum_{n=1}^{3} \alpha_n \exp\left(-\frac{\zeta}{\gamma_n}\right) \right. \\
&\qquad \left. \times \mathbb{E}_{\boldsymbol{G}}\left[\exp\left(-\frac{\left\|\boldsymbol{G}\left(\boldsymbol{A}_i \boldsymbol{p}_{i,j} S_k - \boldsymbol{A}_{\tilde{i}} \boldsymbol{p}_{\tilde{i},\tilde{j}} S_{\tilde{k}}\right)\right\|^2}{\gamma_n}\right)\right]\right] \\
&= \mathbb{E}_{\boldsymbol{H}}\left[\sum_{n=1}^{3} \alpha_n \exp\left(-\frac{\zeta}{\gamma_n}\right) \mathbb{E}_{\boldsymbol{G}}\left[\exp\left(-\frac{\|\boldsymbol{G}\boldsymbol{v}\|^2}{\gamma_n}\right)\right]\right] \\
&= \mathbb{E}_{\boldsymbol{H}}\left[\sum_{n=1}^{3} \alpha_n \exp\left(-\frac{\zeta}{\gamma_n}\right) \left(\frac{\gamma_n}{\|\boldsymbol{v}\|^2 + \gamma_n}\right)^D\right], \quad (23)
\end{aligned}$$

where $\boldsymbol{v} = \boldsymbol{A}_i \boldsymbol{p}_{i,j} S_k - \boldsymbol{A}_{\tilde{i}} \boldsymbol{p}_{\tilde{i},\tilde{j}} S_{\tilde{k}}$, and the last equation is obtained by applying the MGF over the Rayleigh fading channel $\boldsymbol{G}$, which is given in [32].

Remark 1: In the high SNR region, when $\gamma = 1/\sigma^2 \to \infty$ or $\sigma^2 \to 0$, by substituting $\gamma_n = \tau_n \sigma^2$, (23) can be approximated as

$$\mathbb{E}_{\boldsymbol{H},\boldsymbol{G}}[\Delta(\zeta,\xi)] \approx \gamma^{-D} \mathbb{E}_{\boldsymbol{H}}\left[\sum_{n=1}^{3} \frac{\alpha_n {\tau_n}^D}{\|\boldsymbol{v}\|^{2D}} \exp\left(-\frac{\zeta}{\gamma_n}\right)\right]. \quad (24)$$

From (24), we can see that the D diversity antennas are capable of providing D extra orders of receive diversity for performance enhancement.

A. ANALYSIS OF APPROXIMATE ABEP

From (17) – (23), we can learn that the difficulty of deriving a closed-from upper-bound lies in the calculation of the statistical expectation shown in the last line of (23). When M_s is a finite number, the expressions of η_i, $\eta_{\tilde{i}}$, and $\boldsymbol{v}$ are unable to be further simplified, while the distributions of them are all very complicated for further simplification [22]. Hence, in the case of finite M_s, it is extremely hard (if not impossible) for us to derive a closed-form expression for the ABEP upper-bound. Therefore, in this paper, we introduce the Gamma approximation (Gamma-Ap) approach [33] for deriving an approximate APEP upper-bound.

Let us define $\vartheta = \zeta + \xi$. Then, with the aid of the Gamma-Ap, we approximate ϑ as a Gamma distributed random variable with a probability density function (PDF) [34] of

$$f(\vartheta; \lambda_\vartheta, \theta_\vartheta) = \frac{\vartheta^{\lambda_\vartheta - 1}}{\Gamma(\lambda_\vartheta)} {\theta_\vartheta}^{-\lambda_\vartheta} e^{-\vartheta/\theta_\vartheta}, \quad \vartheta \geq 0, \quad (25)$$

where the distribution parameters λ_ϑ and θ_ϑ are given by [33], [34]

$$\begin{aligned}
\lambda_\vartheta &= (\mathbb{E}\{\vartheta\})^2 / \mathbb{E}\left\{(\vartheta - \mathbb{E}\{\vartheta\})^2\right\}, \\
\theta_\vartheta &= \mathbb{E}\left\{(\vartheta - \mathbb{E}\{\vartheta\})^2\right\} / \mathbb{E}\{\vartheta\}. \quad (26)
\end{aligned}$$

From our previous discussion, the analytical solutions to λ_ϑ and θ_ϑ are hard to be mathematically derived. However, the estimates to them can be easily obtained via Monte-Carlo simulations. As declared in [33], only about 10^4 channel realisations are required for achieving sufficiently reliable estimation for λ_ϑ and θ_ϑ. Therefore, in our numerical computations in Section VI, we use 10^4 Monte-Carlo simulations to estimate the distribution parameters of the Gamma PDF in (25). As shown by our results in Section VI, the Gamma-Ap approach is highly effective for the error performance evaluation.

With the aid of the Gamma PDF for ϑ, the expectation at the RHS of (22) can be derived as

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{H},\boldsymbol{G}}[\Delta(\zeta,\xi)] \\
&= \int_0^\infty \Delta(\vartheta) f(\vartheta; \lambda_\vartheta, \theta_\vartheta)\, d\vartheta \\
&= \sum_{n=1}^{3} \int_0^\infty \alpha_n e^{-\vartheta/\gamma_n} \frac{\vartheta^{\lambda_\vartheta - 1}}{\Gamma(\lambda_\vartheta)} (\theta_\vartheta)^{-\lambda_\vartheta} e^{-\vartheta/\theta_\vartheta} d\vartheta \\
&= \sum_{n=1}^{3} \alpha_n \left(\frac{\gamma_n}{\theta_\vartheta + \gamma_n}\right)^{\lambda_\vartheta}, \quad (27)
\end{aligned}$$

where the last equation is obtained due to the formula of (3.326.2) in [35]. Upon substituting (27) into (22), the approximate ABEP upper-bound can be expressed as

$$\begin{aligned}
P_B^{app,ub} = \kappa \sum_{\langle i,j,k\rangle} \sum_{\langle \tilde{i},\tilde{j},\tilde{k}\rangle} D_{\mathrm{H}}\left(\langle i,j,k\rangle, \langle \tilde{i},\tilde{j},\tilde{k}\rangle\right) \\
\times \sum_{n=1}^{3} \alpha_n \left(\frac{\gamma_n}{\theta_\vartheta + \gamma_n}\right)^{\lambda_\vartheta}. \quad (28)
\end{aligned}$$

Let us now turn to analyse the asymptotic ABEP of the JSM systems.

B. ANALYSIS OF ASYMPTOTIC ABEP

Above we have derived the approximate ABEP upper-bound for fixed and relatively small M_s value with the aid of the Gamma-Ap. Alternatively, we can use the asymptotic analysis to derive a closed-from ABEP, by assuming infinite number of antennas. From references [36], [37], and also our results in Section VI, the performance results usually converge fast, and the near asymptotic results can be achieved by a moderate number of antennas.

In the scenario of activating an infinite number of TAs by each transmit pattern, we assume that $M_t > M_s \gg \max\{M_r, D\}$. In this case, we have $\lim_{M_s\to\infty} \boldsymbol{H}_i\boldsymbol{H}_i^H = M_s\boldsymbol{I}_{M_r}$. Therefore, the TZF preprocessing matrix shown in (6) converges to

$$\boldsymbol{P}_i = \eta_i \boldsymbol{H}_i^H \left(\boldsymbol{H}_i\boldsymbol{H}_i^H\right)^{-1} = \boldsymbol{H}_i^H/\sqrt{M_s}, \tag{29}$$

which is in fact the transmitter matched-filtering (TMF) relying preprocessing scheme [38].

Applying the asymptotic result of (29), we can express ζ and ξ in (18) as

$$\zeta = M_s \left\| \boldsymbol{e}_j S_k - \boldsymbol{e}_{\tilde{j}} S_{\tilde{k}} \right\|^2, \tag{30}$$

$$\xi = \left\| \boldsymbol{G}\left(\boldsymbol{A}_i\boldsymbol{H}_i^H\boldsymbol{e}_j S_k - \boldsymbol{A}_{\tilde{i}}\boldsymbol{H}_{\tilde{i}}^H\boldsymbol{e}_{\tilde{j}} S_{\tilde{k}}\right)\right\|^2 / M_s. \tag{31}$$

Accordingly, the expectation of (23) can be expressed as

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{H},\boldsymbol{G}}\left[\Delta\left(\zeta,\xi\right)\right] \\
&= \sum_{n=1}^{3} \alpha_n \exp\left(-\frac{M_s\left\|\boldsymbol{e}_j S_k - \boldsymbol{e}_{\tilde{j}} S_{\tilde{k}}\right\|^2}{\gamma_n}\right) \\
&\quad \times \mathbb{E}_{\boldsymbol{H}}\left[\mathbb{E}_{\boldsymbol{G}}\left[\exp\left(-\frac{\left\|\boldsymbol{G}\left(\boldsymbol{B}_i\boldsymbol{H}^H\boldsymbol{e}_j S_k - \boldsymbol{B}_{\tilde{i}}\boldsymbol{H}^H\boldsymbol{e}_{\tilde{j}} S_{\tilde{k}}\right)\right\|^2}{M_s\gamma_n}\right)\right]\right] \\
&= \sum_{n=1}^{3} \alpha_n \exp\left(-\frac{M_s D_{jk}}{\gamma_n}\right) \mathbb{E}_{\boldsymbol{H}}\left[\mathbb{E}_{\boldsymbol{G}}\left[\exp\left(-\frac{\|\boldsymbol{G}\boldsymbol{\omega}\|^2}{\gamma_n}\right)\right]\right],
\end{aligned} \tag{32}$$

where by definition, $\boldsymbol{B}_i = \boldsymbol{A}_i\boldsymbol{A}_i^H$, $\boldsymbol{B}_{\tilde{i}} = \boldsymbol{A}_{\tilde{i}}\boldsymbol{A}_{\tilde{i}}^H$, $D_{jk} = \|\boldsymbol{e}_j S_k - \boldsymbol{e}_{\tilde{j}} S_{\tilde{k}}\|^2$, and $\boldsymbol{\omega} = (\boldsymbol{B}_i\boldsymbol{H}^H\boldsymbol{e}_j S_k - \boldsymbol{B}_{\tilde{i}}\boldsymbol{H}^H\boldsymbol{e}_{\tilde{j}} S_{\tilde{k}})/\sqrt{M_s}$. Then, using the MGF approach to simplify $\mathbb{E}_{\boldsymbol{G}}[\cdot]$, we simplify (32) to

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{H},\boldsymbol{G}}\left[\Delta\left(\zeta,\xi\right)\right] &= \sum_{n=1}^{3} \alpha_n \exp\left(-\frac{M_s D_{jk}}{\gamma_n}\right) \\
&\quad \times \mathbb{E}_{\boldsymbol{H}}\left[\left(\frac{\gamma_n}{\|\boldsymbol{\omega}\|^2 + \gamma_n}\right)^D\right].
\end{aligned} \tag{33}$$

Finally, substituting (33) into (22), the asymptotic ABEP upper-bound of the JSM system can be formulated as

$$\begin{aligned}
P_B^{asy,ub} &= \kappa \sum_{\langle i,j,k\rangle} \sum_{\langle \tilde{i},\tilde{j},\tilde{k}\rangle} D_{\mathrm{H}}\left(\langle i,j,k\rangle, \langle \tilde{i},\tilde{j},\tilde{k}\rangle\right) \sum_{n=1}^{3} \alpha_n \\
&\quad \times \exp\left(-\frac{M_s D_{jk}}{\gamma_n}\right) \mathbb{E}_{\boldsymbol{H}}\left[\left(\frac{\gamma_n}{\|\boldsymbol{\omega}\|^2 + \gamma_n}\right)^D\right].
\end{aligned} \tag{34}$$

Remark 2: In the high SNR region, rendering $\sigma^2 = 1/\gamma \to 0$, the asymptotic ABEP upper-bound (34) can be approximated as

$$\begin{aligned}
P_B^{asy,ub} &\approx \kappa\gamma^{-D} \sum_{\langle i,j,k\rangle} \sum_{\langle \tilde{i},\tilde{j},\tilde{k}\rangle} D_{\mathrm{H}}\left(\langle i,j,k\rangle, \langle \tilde{i},\tilde{j},\tilde{k}\rangle\right) \\
&\quad \times \mathbb{E}_{\boldsymbol{H}}\left[\|\boldsymbol{\omega}\|^{-2D}\right] \sum_{n=1}^{3} \alpha_n {\tau_n}^D \exp\left(-\frac{M_s D_{jk}}{\gamma_n}\right).
\end{aligned} \tag{35}$$

From (35) we can see that, using D diversity RAs is capable of achieving D orders of diversity for the large-scale JSM systems, which is in accordance with our previous observation stated by *Remark 1*.

After careful investigation, we realised that the expectation term in (34) is hard to be further simplified without imposing approximation. Therefore, we again apply the Gamma-Ap to simplify it. Let $w = \|\boldsymbol{\omega}\|^2$. Then, following the similar procedures as shown in Section V-A, we can approximate w as a Gamma distributed random variable with its parameters expressed as λ_w and θ_w, respectively. In this case, the expectation in (34) can be derived as

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{H}}\left[\left(\frac{\gamma_n}{\|\boldsymbol{\omega}\|^2 + \gamma_n}\right)^D\right] \\
&\approx \int_0^\infty \left(\frac{\gamma_n}{w+\gamma_n}\right)^D \frac{w^{\lambda_w - 1}}{\Gamma\left(\lambda_w\right)\left(\theta_w\right)^{\lambda_w}} e^{-w/\theta_w} dw \\
&\overset{(a)}{=} \frac{\left(\gamma_n\right)^D \left(\theta_w\right)^{-\lambda_w}}{\Gamma\left(\lambda_w\right)} e^{\gamma_n/\theta_w} \int_{\gamma_n}^\infty \frac{\left(\varrho - \gamma_n\right)^{\lambda_w - 1}}{\varrho^D} e^{-\varrho/\theta_w} d\varrho \\
&\overset{(b)}{=} \sqrt{\left(\delta_n\right)^{\lambda_w + D - 1} e^{\delta_n}}\, W_{\frac{-\lambda_w - D + 1}{2}, \frac{-\lambda_w + D}{2}}\left(\delta_n\right).
\end{aligned} \tag{36}$$

Note that in (36), the equality (a) is obtained by applying the variable substitution of $\varrho = w + \gamma_n$, the equality (b) is derived with the aid of the formula of (3.383.4) in [35], $W_{a,b}(z)$ is the Whittaker function defined by the formula of (9.220.4) in [35], and $\delta_n = \gamma_n/\theta_w$. Finally, when substituting (36) into (34), the approximated asymptotic ABEP upper-bound of the JSM system is expressed as

$$\begin{aligned}
P_B^{asy,ub} &\approx \kappa \sum_{\langle i,j,k\rangle} \sum_{\langle \tilde{i},\tilde{j},\tilde{k}\rangle} D_{\mathrm{H}}\left(\langle i,j,k\rangle, \langle \tilde{i},\tilde{j},\tilde{k}\rangle\right) \sum_{n=1}^{3} \alpha_n e^{-\frac{M_s D_{jk}}{\gamma_n}} \\
&\quad \times \sqrt{\left(\delta_n\right)^{\lambda_w + D - 1} e^{\delta_n}}\, W_{\frac{-\lambda_w - D + 1}{2}, \frac{-\lambda_w + D}{2}}\left(\delta_n\right).
\end{aligned} \tag{37}$$

Below, we provide the numerical and simulation results for the performance of the JSM systems.

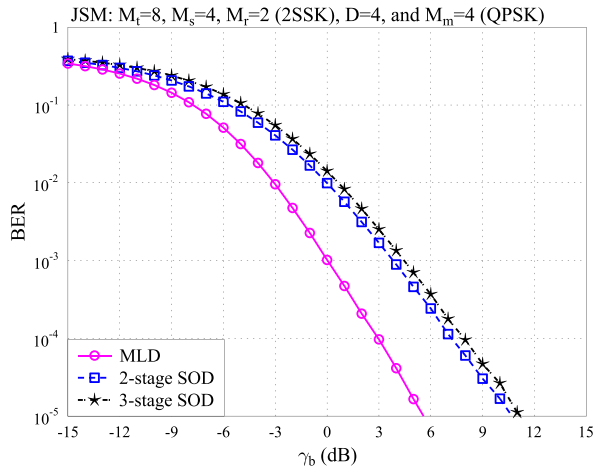


FIGURE 2. BER performance comparison of the JSM systems employing respectively the MLD, two-stage SOD, and the three-stage SOD, when communicating over Rayleigh fading channels.

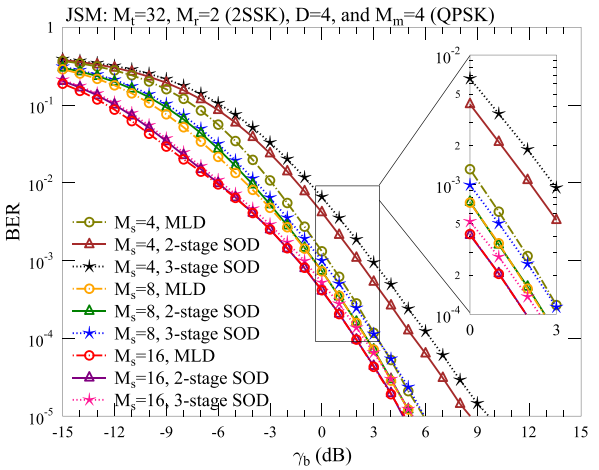


FIGURE 3. BER performance of the JSM systems with the MLD, two-stage SOD, and the three-stage SOD, when communicating over Rayleigh fading channels.

VI. PERFORMANCE RESULTS AND DISCUSSION

In this section, we demonstrate and compare the bit error rate (BER) performance of our proposed JSM systems with the MLD, two-stage SOD, and the three-stage SOD, respectively. Both numerical results evaluated based on formulas and simulation results are provided. Note that in all figures, the common system parameters are detailed on the top of the figures, where NSSK stands for $M_r = N$, and γ_b denotes the average SNR per bit.

First, Fig. 2 compares the BER versus average SNR per bit performance of the JSM systems employing the MLD, two-stage SOD and three-stage SOD. Explicitly, the MLD achieves the best BER performance among the three detectors. In comparison with the MLD, the two-stage SOD results in a remarkable performance loss, requiring about an extra 5 dB SNR to achieve the same BER of 10^{-4} as the MLD. While benefiting a complexity reduction, the three-stage SOD also experiences some performance loss in comparison with the two-stage SOD. Nevertheless, the performance's difference between the two SODs is marginal, and is within 1 dB over the whole SNR range considered.

Second, in Figs. 3-6, we characterise the impact of the system parameters of M_s, M_r, M_m and D on the error performance of the JSM systems, when the MLD, two-stage SOD and the three-stage SOD are respectively employed. In Figs. 3-5, we demonstrate the effect of the number of bits per symbol on the BER performance, via changing the number of bits delivered by the TSM, RSM or the APM symbol. Specifically, in Fig. 3, M_s is increased from 4 to 8 and then to 16, which corresponds to $b_J = 6$, 5 or 4. From the results of Fig. 3, we can have the following observations. First, the BER performance of the JSM systems improves as M_s increases. This is due to the fact that the total number of bits per symbol is reduced, and that the transmit diversity increases, as M_s increases. Second, for the JSM systems employing the MLD, as M_s increases, the BER improvement in the low SNR region is more declared than that in the medium to high SNR regions. The reason behind this can be explained as follows. In the low SNR region, the performance of the JSM systems is dominated by the TSM. Therefore, as M_s increases, significant performance improvement is observed due to the transmit diversity. When the SNR increases, resulting in that the TSM becomes sufficiently reliable, the BER performance of the JSM system is then jointly determined by the TSM, RSM and the APM. Correspondingly, the BER improvement due to the increase of M_s become less significant. Third, for the JSM systems with the two SOD schemes, the BER performance improves significantly as M_s is increased from 4 to 8, while only slight improvement is observed, when M_s is increased from 8 to 16. Finally, as shown in Fig. 3, with the increase of M_s, the BER achieved by the two SODs becomes more and more close to the BER attained by the MLD. This is because the transmit diversity gain achieved by the JSM systems increases, as the value of M_s is increased from 4 to 8 and then to 16, resulting in that the TSM symbol detected by the two-stage and three-stage SODs becomes more reliable.

In Fig. 4, we depict the BER performance of the JSM systems, when M_r for RSM is increased from 2 to 4 and then to 8, which result in that the total number of bits per JSM symbol $b_J = 4, 5$ and 6. From Fig. 4, we can explicitly see that the performance gap between the SODs and the MLD becomes bigger as M_r increases. In general, for the three types of detectors considered, the BER performance becomes worse, as M_r increases, i.e., as the number of bits per RSM symbol increases. This is because, for a given M_s, the transmit diversity decreases as M_r increases. As seen in Fig. 4, for the MLD, the error performance of the JSM systems with 4SSK is slightly better than that of the JSM systems with 2SSK in the high SNR region. This is because the 4SSK is capable of enhancing more the QPSK's BER performance than the 2SSK is, and the corresponding performance enhancement can compensate for the performance loss caused by the decrease of the transmit diversity, when M_r is increased from 2 to 4. As demonstrated in Fig. 4, the error performance of the two SODs is very close to each other, which implies that the APM symbol can usually be reliably detected, once the

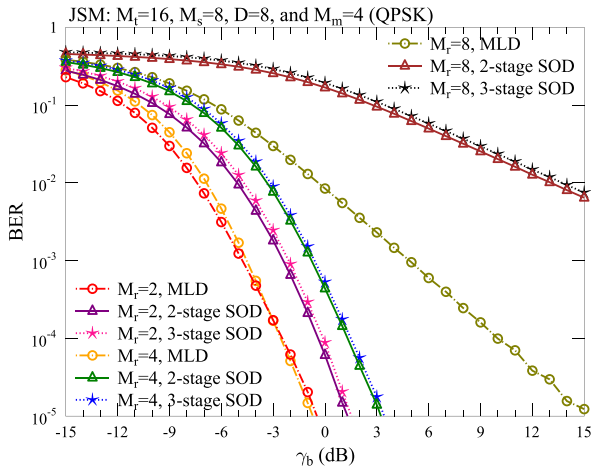


FIGURE 4. BER performance of the JSM systems with the MLD, two-stage SOD, and the three-stage SOD, when communicating over Rayleigh fading channels.

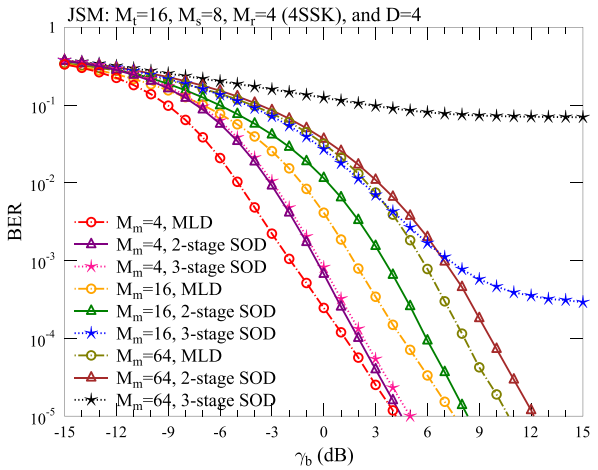


FIGURE 5. BER performance of the JSM systems with the MLD, two-stage SOD, and the three-stage SOD, when communicating over Rayleigh fading channels.

RSM symbol is correctly detected. Furthermore, Fig. 4 shows that, when the two-stage or three-stage SOD is employed, the performance loss generated by the increase of M_r from 4 to 8 is much more significant than that resulted from the increase of M_r from 2 to 4. The reason behind is that, when given $M_s = 8$, there is still sufficient transmit diversity for detection of the RSM symbol, when $M_r = 4$. However, in the case of $M_r = 8$, there is no transmit diversity anymore, hence making the detection of the RSM symbol less reliable.

In Fig. 5, we characterise the impact of M_m on the BER performance of the JSM systems, by increasing M_m from 4 to 16 and to 64, resulting that $b_J = 5$, 7 and 9. Observing Fig. 5, we can find that the error performance of the JSM systems employing any type of detector degrades significantly, as M_m increases. Furthermore, the performance loss of the three-stage SOD against that of the MLD or the two-stage SOD becomes more significant, as M_m increases, while the loss of the two-stage SOD against that of the MLD is marginal, when M_m is increased from 4 to 16 and to 64. The above observations occur due to the fact that, when M_m is relative large, the system's BER performance is dominated by the detection of the APM symbol. Additionally, when the three-stage SOD is employed, we observe that the BER performance curves appear error floors in the high SNR region, when the 16QAM or 64QAM is used. This observation is due to the detection errors introduced during the second stage detection, when the modulation scheme of 16QAM or 64QAM is employed.

In Fig. 6, we demonstrate the BER performance of the JSM systems employing the three types of detectors, when the number of diversity antennas D equals 1, 2, and 4, respectively. Explicitly, the error performance of the JSM systems improves as D increases. For the JSM systems with a given number of diversity antennas, the MLD achieves the best BER performance, while the performance achieved by the two-stage SOD converges to that of the MLD as SNR increases. This is because the detection of the RSM symbol

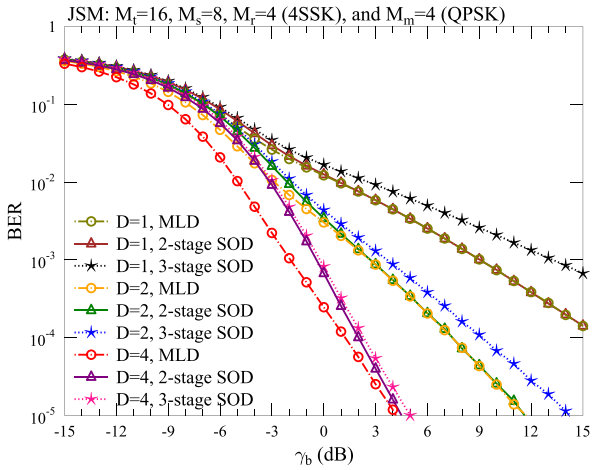


FIGURE 6. BER performance of the JSM systems employing the MLD, two-stage and three-stage SODs, when communicating over Rayleigh fading channels.

becomes highly reliable, when the SNR is sufficiently high, making the overall BER performance mainly dominated by the joint detection of the TSM and APM symbols. When comparing the BER performance of the two SODs, we can see that the performance difference between them becomes smaller and smaller, as D is increased from 1 to 2 and then to 4. This observation explains that using more diversity antennas is beneficial to the TSM symbol's detection in the three-stage SOD, resulting in that the joint detection in the two-stage SOD and separate detection in the three-stage SOD achieve similar performance.

From Figs. 3-6, we are implied that once M_t, $(M_r + D)$ and b_J are given, there exists an optimal selection for b_t, b_r and b_m, which results in the best BER performance. To demonstrate this, in Fig. 7, we let $M_t = 32$, $M_r + D = 10$, and $b_J = 6$, and depict the BER performance results of the JSM systems with various selections of b_t, b_r and b_m, when utilising the MLD detection. In order to achieve the different combinations of b_t, b_r and b_m, we set $M_s \in \{4, 8, 16\}$,

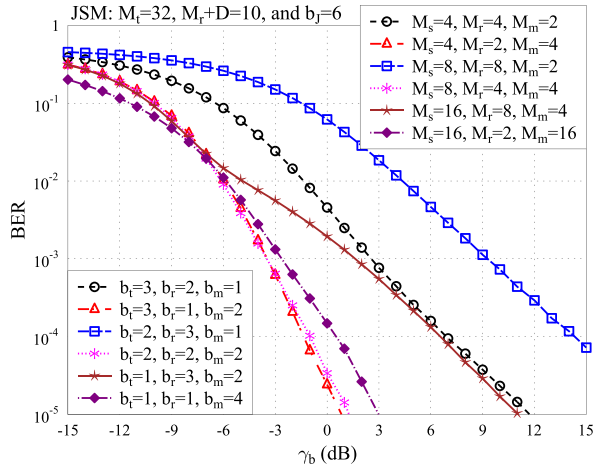


FIGURE 7. Comparison of the BER performance of the JSM systems employing MLD, when communicating over Rayleigh fading channels.

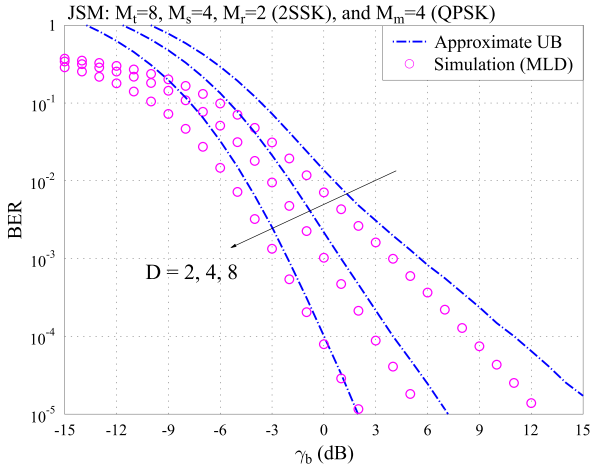


FIGURE 8. Comparison between the simulated BER and the approximate ABEP upper-bound (UB) for the JSM systems with MLD, when communicating over Rayleigh fading channels.

$M_r \in \{2, 4, 8\}$, and $M_m \in \{2, 4, 16\}$. For the clarity of presentation, the different selections of b_t, b_r and b_m, and the corresponding M_s, M_r, and M_m values are all detailed in Fig. 7.

From the results shown in Fig. 7, we can see that the best BER performance is achieved, when the JSM system is configured as $b_t = 3$, $b_r = 1$ and $b_m = 2$. In this case, the TSM symbols, whose detection reliability dominates the detection performance of the JSM system, can be more reliably detected than in the other cases. Moreover, when increasing the number of bits delivered by the TSM symbol, the performance loss given by the RSM and APM detection is marginal, when compared with the other cases. By comparing the BER performance of the JSM systems having the same b_t, we can observe that a smaller value of b_r yields a better BER performance. This is because both the transmit and receive diversity increases, as b_r, i.e., M_r decreases, resulting in more reliable joint detection of the TSM, RSM, and APM symbols.

Having demonstrated the BER performance obtained by simulations for the JSM systems with various detection schemes, let us now show the approximate ABEP upper-bound (UB) of (28) in Fig. 8 and the asymptotic ABEP-UB of (37) in Fig. 9. Both of them are accompanied by the simulation results for their validation. As shown in Fig. 8, we set $M_t = 8$, $M_s = 4$ for the regular-scale JSM system, while in Fig. 9, we set $M_t = 128$, $M_s = 64$ for the large-scale JSM system. The other parameters for system configurations are detailed in the figures.

From the results shown in Figs. 8 and 9, we can attain the following observations. First, for both the regular-scale and large-scale JSM systems, as shown by the simulated BER and the corresponding UB (approximate/asymptotic), the BER performance becomes better, as D increases. Straightforwardly, this is because the receive diversity gain becomes larger as D increases. Second, although the asymptotic UBs are witnessed to be tighter than the approximate UBs in the low SNR region, the approximate and asymptotic UBs

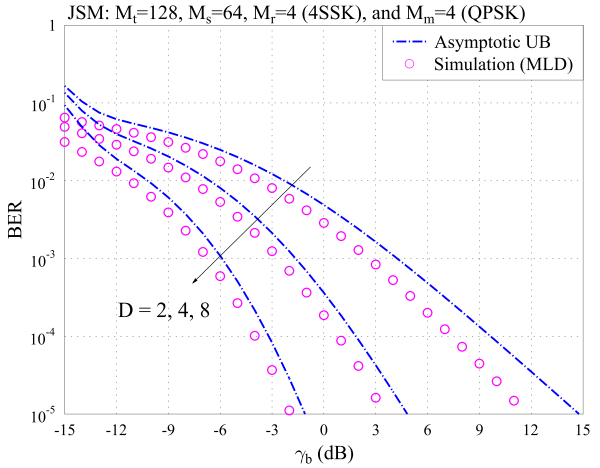


FIGURE 9. Comparison between the simulated BER and the asymptotic ABEP upper-bound (UB) for the JSM systems with MLD, when communicating over Rayleigh fading channels.

for all the cases are in general reasonably tight. Hence, we can be convinced that both the MGF and the Gamma-Ap approaches are effective for deriving the approximate ABEP-UB of (28) and the asymptotic ABEP-UB of (37). Third, with the increase of D, both the approximate ABEP-UB and the asymptotic ABEP-UB become tighter. The reason behind this is that, when D increases, more independent Gaussian-noise like observations are generated, making the approximation implemented by utilising the Gamma-Ap approach more effective.

Based on the above observations, we can be implied that both the approximate ABEP-UB and the asymptotic ABEP-UB derived in Section V can be applied to predict the BER performance of the JSM systems, provided that the SNR is sufficiently high, resulting in that the BER is about or below 10^{-2}.

VII. SUMMARY AND CONCLUSIONS

We proposed and studied a JSM system, which transmits information via exploiting the degrees of freedom jointly provided by the indices of transmit patterns, indices of RAs, and

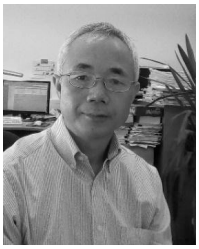
the conventional APM. In the proposed JSM systems, extra RAs were utilised to provide receive diversity. Hence, when properly designed, the JSM system is capable of achieving transmit diversity, receive diversity, and multiplexing gain. In this paper, the transmitter preprocessing based on TZF was employed for principle description and performance analysis. Explicitly, other types of preprocessing schemes can be employed. In order to facilitate the implementation of the JSM systems at different complexity levels, we proposed and investigated two types of detectors, namely the joint MLD and SOD. Furthermore, two SODs were proposed, which are the two-stage SOD and three-stage SOD. To gain insight into the system performance, we analysed the approximate and asymptotic ABEP of the JSM systems employing MLD, and correspondingly the analytical results were derived, when both the regular- and large-scale JSM systems were assumed. Our simulation results demonstrated that the MLD is capable of achieving the best error performance among the three detection schemes. However, the performance achieved by the two SODs converges to that of the MLD as SNR increases. Moreover, in most cases of practically interested, the performance loss of the three-stage SOD against that of the two-stage SOD is marginal. Finally, by conducting the comparisons between the simulated BER and numerical ABEP results, the analytical expressions derived in Section V were validated. Our studies in this paper demonstrate that the JSM is highly flexible for implementation and online configuration, according to the communications environments as well as the constraints on performance and implementation complexity.

REFERENCES

[1] M. Di Renzo, H. Haas, and P. M. Grant, "Spatial modulation for multiple-antenna wireless systems: A survey," *IEEE Commun. Mag.*, vol. 49, no. 12, pp. 182–191, Dec. 2011.
[2] P. Yang, M. Di Renzo, Y. Xiao, S. Li, and L. Hanzo, "Design guidelines for spatial modulation," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 6–26, 1st Quart., 2015.
[3] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.
[4] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.
[5] J. Jeganathan, A. Ghrayeb, L. Szczecinski, and A. Ceron, "Space shift keying modulation for MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3692–3703, Jul. 2009.
[6] M. Di Renzo and H. Haas, "A general framework for performance analysis of space shift keying (SSK) modulation for MISO correlated Nakagami-*m* fading channels," *IEEE Trans. Commun.*, vol. 58, no. 9, pp. 2590–2603, Sep. 2010.
[7] L.-L. Yang, "Signal detection in antenna-hopping space-division multiple-access systems with space-shift keying modulation," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 351–366, Jan. 2012.
[8] M. Wen, X. Cheng, H. V. Poor, and B. Jiao, "Use of SSK modulation in two-way amplify-and-forward relaying," *IEEE Trans. Veh. Technol.*, vol. 63, no. 3, pp. 1498–1504, Mar. 2014.
[9] R. Mesleh, H. Haas, C. W. Ahn, and S. Yun, "Spatial modulation—A new low complexity spectral efficiency enhancing technique," in *Proc. Chinacom*, Oct. 2006, pp. 1–5.
[10] R. Y. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228–2241, Jul. 2008.
[11] Y. Yang and B. Jiao, "Information-guided channel-hopping for high data rate wireless communication," *IEEE Commun. Lett.*, vol. 12, no. 4, pp. 225–227, Apr. 2008.
[12] R. Mesleh, M. Di Renzo, H. Haas, and P. M. Grant, "Trellis coded spatial modulation," *IEEE Trans. Wireless Commun.*, vol. 9, no. 7, pp. 2349–2361, Jul. 2010.
[13] E. Basar, U. Aygolu, E. Panayirci, and H. V. Poor, "Space-time block coded spatial modulation," *IEEE Trans. Commun.*, vol. 59, no. 3, pp. 823–832, Mar. 2011.
[14] P. Yang, Y. Xiao, Y. Yu, and S. Li, "Adaptive spatial modulation for wireless MIMO transmission systems," *IEEE Commun. Lett.*, vol. 15, no. 6, pp. 602–604, Jun. 2011.
[15] M. Di Renzo and H. Haas, "Bit error probability of SM-MIMO over generalized fading channels," *IEEE Trans. Veh. Technol.*, vol. 61, no. 3, pp. 1124–1144, Mar. 2012.
[16] J. Wang, S. Jia, and J. Song, "Generalised spatial modulation system with multiple active transmit antennas and low complexity detection scheme," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1605–1615, Apr. 2012.
[17] R. Mesleh, S. S. Ikki, and H. M. Aggoune, "Quadrature spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 64, no. 6, pp. 2738–2742, Jun. 2015.
[18] Y. Bian, X. Cheng, M. Wen, L. Yang, H. V. Poor, and B. Jiao, "Differential spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 64, no. 7, pp. 3262–3268, Jul. 2015.
[19] A. Stavridis, D. Basnayaka, S. Sinanovic, M. Di Renzo, and H. Haas, "A virtual MIMO dual-hop architecture based on hybrid spatial modulation," *IEEE Trans. Wireless Commun.*, vol. 62, no. 9, pp. 3161–3179, Sep. 2014.
[20] L.-L. Yang, "Transmitter preprocessing aided spatial modulation for multiple-input multiple-output systems," in *Proc. IEEE 73rd VTC (Spring)*, May 2011, pp. 1–5.
[21] R. Zhang, L.-L. Yang, and L. Hanzo, "Generalised pre-coding aided spatial modulation," *IEEE Trans. Wireless Commun.*, vol. 12, no. 11, pp. 5434–5443, Nov. 2013.
[22] R. Zhang, L.-L. Yang, and L. Hanzo, "Error probability and capacity analysis of generalised pre-coding aided spatial modulation," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 364–375, Jan. 2015.
[23] R. Zhang, L.-L. Yang, and L. Hanzo, "Performance analysis of non-linear generalized pre-coding aided spatial modulation," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6731–6741, Oct. 2016.
[24] J. Zheng, "Fast receive antenna subset selection for pre-coding aided spatial modulation," *IEEE Wireless Commun. Lett.*, vol. 4, no. 3, pp. 317–320, Jun. 2015.
[25] M. Zhang, M. Wen, X. Cheng, and L. Yang, "Pre-coding aided differential spatial modulation," in *Proc. IEEE GLOBECOM*, Dec. 2015, pp. 1–6.
[26] M. Zhang, M. Wen, X. Cheng, and L. Yang, "A dual-hop virtual MIMO architecture based on hybrid differential spatial modulation," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6356–6370, Sep. 2016.
[27] R. Pizzio, B. F. Uchôa-Filho, M. Di Renzo, and D. L. Ruyet, "Generalized spatial modulation for downlink multiuser MIMO systems with multicast," in *Proc. IEEE 27th PIMRC*, Sep. 2016, pp. 1–6.
[28] J. Li, M. Wen, M. Zhang, and X. Cheng, "Virtual spatial modulation," *IEEE Access*, vol. 4, pp. 6929–6938, Nov. 2016.
[29] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels*, 2nd ed. New York, NY, USA: Wiley, 2005.
[30] M. I. Irshid and I. S. Salous, "Bit error probability for coherent M-ary PSK systems," *IEEE Trans. Commun.*, vol. 39, no. 3, pp. 349–352, Mar. 1991.
[31] M. Chiani, D. Dardari, and M. K. Simon, "New exponential bounds and approximations for the computation of error probability in fading channels," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 840–845, Jul. 2003.
[32] 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.
[33] J. Shi, C. Dong, and L.-L. Yang, "Performance comparison of cooperative relay links with different relay processing strategies: Nakagami/Gamma approximation approaches," *EURASIP J. Wireless Commun. Netw.*, vol. 53, pp. 1–17, Apr. 2014.
[34] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed. New York, NY, USA: McGraw-Hill, 2002.
[35] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. San Diego, CA, USA: Academic, 2007.
[36] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

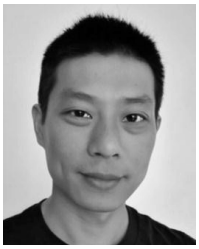
[37] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.
[38] L.-L. Yang, *Multicarrier Communications*. Hoboken, NJ, USA: Wiley, 2009.

CHAOWEN LIU received the B.S. degree in electrical engineering from the Henan University of Science and Technology, Luoyang, China, in 2011. He is currently working toward the Ph.D. degree at the Institute of Information Engineering, Xi'an Jiaotong University, Xi'an, China. From 2015 to 2017, he was a visiting Ph.D. student with Prof. L.-L. Yang at the School of Electronics and Computer Science, University of Southampton, U.K. His main research interests include spatial-domain modulation techniques, physical layer security, full-duplex cooperative wireless communications, index modulation for 5G wireless networks, and positioning in wireless sensor networks.



LIE-LIANG YANG (M'98–SM'02–F'16) received the B.Eng. degree in communications engineering from Tongji University, Shanghai, China, in 1988 and the M.Eng. and Ph.D. degrees in communications and electronics from Northern (Beijing) Jiaotong University, Beijing, China, in 1991 and 1997, respectively. From 1997 to 1997, he was a Visiting Scientist at the Institute of Radio Engineering and Electronics, Czech Academy of Sciences. Since 1997, he has been with the University of Southampton, U.K., where he is currently a Professor of Wireless Communications at the School of Electronics and Computer Science. He has published 350 research papers in journals and conference proceedings, authored/coauthored three books, and also published several book chapters. His research interests in wireless communications, wireless networks, and signal processing for wireless communications, molecular communications, and nano-networks. He is a Fellow of IET (previously IEE) in the U.K., and he is a Distinguished Lecturer of the IEEE Vehicular Technology Society. He has served as an Associate Editor for several academic journals, co-organized several special issues, and acted as different roles for conference organization.

WENJIE WANG (M'10) received the 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 at the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE, USA. He is currently 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.



FASONG WANG received the B.Sc. and M.Sc. degrees in applied mathematics from China University of Geosciences, China, in 2002 and 2005, respectively, and the Ph.D. degree in communication engineering from Xidian University in 2013. He was a Senior Engineer at the 27th Research Institute, China Electronics Technology Group Corporation, Zhengzhou, China, from 2005 until 2010. He is currently an Associate Professor at the School of Information Engineering, Zhengzhou University, China. He is also a Visiting Scholar at the University of Southampton, U.K. His current research interests include secrecy communication, blind signal processing, compressed sensing, and sparse representation and their applications. He has published over 40 journals and conference papers on these topics.