Differential STBC-SM Scheme for Uplink Multi-user Massive MIMO Communications: System Design and Performance Analysis

Jing Zhu, Lixia Xiao, *Member, IEEE*, Pei Xiao, *Senior Member, IEEE*, Atta Quddus, Chang He, and Lei Lu

Abstract—In this paper, a novel differential space-time block coded spatial modulation (differential STBC-SM) is proposed for uplink multi-user massive multiple-input multiple-output (MI-MO) communications, which combines the concept of differential coding and STBC-SM to enhance the diversity benefits in the absence of the channel state information (CSI). The transmission structure of the proposed system is on a block basis, where each block contains two sub-blocks. More specifically, the first sub-block only conveys amplitude and phase modulation (APM) symbol bits, since its transmit antennas (TAs) obey a pre-designed activation pattern, which do not carry any information bit. For the second sub-block, the input bits are modulated to STBC-SM matrices, which are then differentially coded between two adjacent sub-blocks. Moreover, a novel block-by-block based non-coherent detector is presented. Finally, we derive an upper bound on the average bit error probability (ABEP) by using the moment generating function (MGF). Our simulation results show that the proposed differential STBC-SM transmission structure is able to acquire considerable bit error rate (BER) performance improvements compared to both the conventional differential spatial modulation (DSM) and differential Alamouti schemes.

Index Terms—Massive MIMO, differential STBC-SM, non-coherent detection, average bit error probability, multi-user uplink communications.

I. INTRODUCTION

MASSIVE multiple-input multiple-output (MIMO) transmission technique is regarded as an important technology for future 5G communication systems since it significantly improve throughput and energy efficiency compared to conventional MIMO systems [1]-[4]. A large number of antennas are employed in massive MIMO systems, which also requires a large number of radio frequency (RF) chains to be connected with antennas, resulting in high hardware cost, energy consumption and signal processing complexity. Under this circumstance, spatial modulation (SM) schemes have been developed as a new MIMO architecture equipped with only

one RF chain, where the information bits are transmitted by spatial symbols (i.e., the activated antenna indices) as well as digital symbols (i.e., the amplitude and phase modulation (APM) symbols) [5]-[7]. Combining massive MIMO and SM (massive SM-MIMO), superior bit error rate (BER), high spectral efficiency and energy efficiency can be simultaneously achieved [8]-[10].

In massive SM-MIMO, one of the major challenges is channel estimation, because the exact channel state information (C-SI) is essential to obtain a satisfactory performance. However, the huge computational complexity and high signalling-to-data ratio brought by channel estimation represents a bottleneck problem to prevent its practical deployment. The situation is further exacerbated in high mobility scenarios where excessive pilot overhead is incurred due to short coherence time. To reduce the computational and pilot overhead, and achieve better spectrum efficiency, differential spatial modulation (DSM) has been proposed in [11]-[13], which utilizes a single RF transmitter structure without the knowledge of CSI at the receiver. Because of the CSI-free transmission structure, the DSM scheme is deemed to a prospective massive MIMO transmission program for the future wireless communications.

Authors in [11] first proposed the concept of DSM in 2013, which utilizes the time and spatial degrees of freedom to facilitate the differential modulation. The performance analysis for the two transmit antennas (TAs) based DSM was presented in [12]. Full square antenna matrix based square DSM (FSAM-SDSM) schemes have been developed in [13]-[19], in which N_t (N_t is the number of the TAs) amplitude phase modulation (APM) symbols are transmitted by one of the activated $N_t \times N_t$ antenna matrices. However, in the FSAM-SDSM transmission scheme, Q ($Q = 2^{\lfloor \log_2(N_t!) \rfloor}$) antenna matrices (AMs) are legitimate, resulting in the challenging issue of bit-to-symbol mapping and demapping since the value of Q soars as N_t increases. To avoid this issue, a novel square DSM structure, dubbed full diversity based square DSM (FD-SDSM), was proposed in [20]-[21], where only $Q = N_t$ AMs are used to simplify the bit-to-symbol mapping. To be specific, in the FD-SDSM system, only a single APM symbol is transmitted through an activated AM via N_t time slots, which indicates that the transmission rate of FD-SDSM is obviously lower than that of its coherent SM counterpart. It can thus be concluded that the existing SDSM schemes are not suitable for downlink massive MIMO scenarios.

To address this limitation, authors in [22] proposed a rectangular DSM (RDSM) scheme for downlink communications.

In RDSM scheme, the antenna matrix is a rectangular matrix, so the length of the RDSM transmission symbol can be significantly reduced, which is more suitable for massive MIMO communications. Besides, [23] presented STBC aided RDSM (STBC-RDSM) scheme for massive MIMO downlink communications, which combines RDSM and STBC to reap their respective advantages. The authors in [24] proposed a simplified RDSM scheme and assigned a new design metric for forgetting factor to mitigate error propagation.

To the best of our knowledge, there are only two papers in the literature focused on the application of DSM in uplink massive MIMO communications. In [25], an energy detection based DSM scheme was presented for single user uplink massive MIMO communications. [26] extended DSM scheme to multi-user scenarios and presented a new non-coherent detection to avoid the error propagation. However, in uplink communications, the aforementioned DSM schemes only utilize the multiplexing gain of multiple TAs without the knowledge of CSI, these transmission schemes have not exploited the potential transmit diversity of MIMO systems. Inspired by the STBC-RDSM transmission structure [23] and multi-user DSM design for uplink communications [26], we propose a novel transmission scheme, namely differential STBC-SM, for uplink multi-user massive MIMO communications to achieve a significant diversity gain.

Our major contributions of this paper can be summarized as follows.

- For uplink multi-user massive MIMO communications, we propose a general differential STBC-SM transmission framework which exploits the synergistic amalgamation of differential coding and STBC-SM to further enhance diversity gain without the knowledge of CSI compared to the existing transmission mechanisms.
- We propose a novel non-coherent detector for our proposed differential STBC-SM system, which eases the impact of inter-user interference while preventing error propagation. The proposed detector is implemented block-by-block. Specifically, the detection of the first sub-block is the same as the conventional single user non-coherent detection because each column of the transmission matrices in the first sub-block is mutually orthogonal. As for the second sub-block, the transmit antenna combinations for each user are first estimated based on the correlation between the received signals of the two sub-blocks, then the APM symbol pairs can be estimated exploiting the selected antenna combination.
- We provide the derivation of the BER performance based on non-coherent detector. In particular, a closed-form expression of the pair-wise error probability (PEP) for the proposed scheme is derived, the accuracy of the analysis is verified by our simulation results.

The rest of this paper is organized as follows. In Section II, an overview of the application of DSM in uplink multi-user massive MIMO communications is described. Section III introduced the transmission structure for the proposed differential STBC-SM scheme. This is followed by the proposed non-coherent detector in Section IV. The BER performance

of the proposed differential STBC-SM scheme is derived in Section V. Section VI provides simulation and performance comparison results. Finally, the conclusion of this paper is presented in Section VII.

Notations: $|\cdot|$, $\binom{n}{k}$, $n!$ and $\lfloor \cdot \rfloor$ refer to the absolute value, the binomial coefficient, factorial and the floor operation, respectively. $(\cdot)^H$, $(\cdot)^*$ and $(\cdot)^{-1}$ denote Hermitian transpose, conjugate and inverse, respectively. $\mathrm{diag}\,(\cdot)$ and $E\,(\cdot)$ stand for the diagonal and expectation operation, respectively. $\otimes$ denotes the Kronecker product operation between two matrices and $\mathrm{Re}\,\{\cdot\}$ represents the real part of a complex variable. $\|\mathbf{A}\|_F$ denotes the Frobenius norm of $\mathbf{A}$, while $\mathbf{I}_N$ stands for an N-dimensional identity matrix. $\mathbb{C}^{M \times N}$ represents the space of $M \times N$ complex matrices. Besides, $\mathrm{supp}(\mathbf{x})$ denotes the index of non-zero element position of vector $\mathbf{x}$ and $\mathrm{mod}\,(a,b)$ represents the remainder of a/b.

II. AN OVERVIEW OF DSM IN UPLINK MULTI-USER MASSIVE MIMO COMMUNICATIONS

Firstly, we will briefly explain why conventional DSM scheme is not applicable to uplink multi-user scenarios. Without the loss of generality, we use a two-user communication system as an example in order to ease understanding. In this case, the DSM transmission matrix for each user can be expressed as

$$\mathbf{S}_1(b) = \mathbf{S}_1(b-1)\mathbf{X}_1(b), \qquad (1)$$

and

$$\mathbf{S}_2(b) = \mathbf{S}_2(b-1)\mathbf{X}_2(b). \qquad (2)$$

Hence, the received signal can be formulated as

$$\begin{aligned}
\mathbf{Y}(b) &= \mathbf{H}(b)\mathbf{S}(b) + \mathbf{V}(b) \\
&= \begin{bmatrix} \mathbf{H}_1(b) & \mathbf{H}_2(b) \end{bmatrix} \begin{bmatrix} \mathbf{S}_1(b) \\ \mathbf{S}_2(b) \end{bmatrix} + \mathbf{V}(b).
\end{aligned} \qquad (3)$$

From Eq. (3), we can see that it is impossible to establish a relationship between $\mathbf{Y}(b)$, $\mathbf{Y}(b-1)$, $\mathbf{X}_1(b)$ and $\mathbf{X}_2(b)$ which are independent of both $\mathbf{H}_1(b)$ and $\mathbf{H}_2(b)$. In other words, we can estimate $\hat{\mathbf{X}}(b)$ using the energy detector, but it is a superposition of $\mathbf{X}_1(b)$ and $\mathbf{X}_2(b)$, which can not be distinguished at the receiver. As such, the conventional DSM is not suitable for the uplink multi-user transmission systems.

To address the aforementioned issue, the authors in [26] presented a new block based DSM transmission structure to extend DSM scheme to multi-user scenarios. Consider an uplink transmission scenario consisting of K users, where each user has N_t TAs, and a base station (BS) equips with numerous receive antennas (RAs) N_r $(N_r \gg N_t)$. The transmission of the multi-user DSM (MU-DSM) scheme in [26] is carried out on the basis of block, which involves two sub-blocks with the lengths of KN_t and L^1, respectively. Note that, in the rest of this paper, the superscript (a) and subscript k represent the a-th sub-block ($a = 1$ or 2) and the k-th user ($k = 1,...,K$), respectively. Besides, the t-th time slot in the b-th block is presented by the ordered pair (b,t).

---

[1] L is a designed parameter based on the coherence time of the channel T_c. To be specific, the total length of a transmission block is no more than T_c, i.e., $KN_t + L \leq T_c$.

Mapping rule for the first sub-block: The TAs of all users in the first sub-block are activated according to a pre-designed TA pattern. To be specific, user k activates its 0-th antenna to $(N_t - 1)$-th antenna during N_t consecutive time slots, and all other users keep silent at the same time. Thus, the activated TA index, $q_k^{(1)}(b,t)$, is equal to $\text{mod}(t, N_t)$, $t = (k-1)N_t, ..., kN_t - 1$. Subsequently, $\mathcal{B}^{(1)} = KN_t\log_2(M)$ bits are modulated to KN_t M-PSK symbols $s_k^{(1)}(b,t), t \in [0, KN_t - 1]$ that are transmitted by the pre-designed TA pattern. As a result, the DSM vector for user k in the first sub-block, $\mathbf{s}_k^{(1)}(b,t) \in \mathbb{C}^{N_t \times 1}$, can be expressed as

$$\mathbf{s}_k^{(1)}(b,t) = \begin{cases} \mathbf{s}_k^{(1)}(b-1,t)s_k^{(1)}(b,t), & t \in [(k-1)N_t, kN_t-1] \\ \mathbf{0}, & \text{otherwise} \end{cases}$$

(4)

where $\mathbf{s}_k^{(1)}(b-1,t)$ is the DSM vector for user k in the $(b-1)$-th transmission block. Since there is only one activated TA in each time slot, $\mathbf{s}_k^{(1)}(b,t)$ includes only one non-zero element whose position index is $q_k^{(1)}(b,t)$. In addition, $q_k^{(1)}(b,t)$ is defined by

$$q_k^{(1)}(b,t) = \text{supp}(\mathbf{s}_k^{(1)}(b,t)), t \in [(k-1)N_t, kN_t - 1]. \quad (5)$$

Mapping rule for the second sub-block: All K users transfer information at the same time in the second sub-block. $\mathcal{B}^{(2)} = KL(\log_2(N_t)+\log_2(M))$ bits are transmitted by the activated TA index as well as M-PSK symbol. Specifically, in the t-th time slot, $(\log_2(N_t)+\log_2(M))$ bits transmitted by user k are separated into two groups: 1) $\log_2(N_t)$ bits are first mapped to the index of the activated antenna $q_k^{(2)}(b,t)$; and 2) $\log_2(M)$ bits are modulated to a M-PSK symbol $s_k^{(2)}(b,t)$ transmitted by the activated antenna. Thus, the transmitted DSM vector is presented as

$$\mathbf{s}_k^{(2)}(b,t) = \mathbf{s}_k^{(1)}(b,\tilde{t})s_k^{(2)}(b,t),$$
$$t \in [KN_t, KN_t + L - 1], \tilde{t} \in [(k-1)N_t, kN_t - 1], \quad (6)$$

where $\tilde{t}$ is the index of time slot in the first sub-block where $q_k^{(1)}(b,\tilde{t}) = q_k^{(2)}(b,t)$.

Let us denote $\mathbf{H}_k^{(1)}(b,t)$ and $\mathbf{H}_k^{(2)}(b,t) \in \mathbb{C}^{N_r \times N_t}$ as the unknown fading channel matrices between user k and the BS of the first and second sub-blocks, respectively. Moreover, $\mathbf{v}_k^{(1)}(b,t)$ and $\mathbf{v}_k^{(2)}(b,t) \sim \mathcal{CN}(\mathbf{0}, \sigma_v^2\mathbf{I}_{N_r})$ are the additive white Gaussian noise (AWGN) matrices. The received signals $\mathbf{y}^{(1)}(b,t)$ and $\mathbf{y}^{(2)}(b,t)$ are presented as

$$\mathbf{y}^{(1)}(b,t) = \sum_{k=1}^K \mathbf{H}_k^{(1)}(b,t)\mathbf{s}_k^{(1)}(b,t) + \mathbf{v}^{(1)}(b,t),$$
$$t \in [0, KN_t - 1], \quad (7)$$

and

$$\mathbf{y}^{(2)}(b,t) = \sum_{k=1}^K \mathbf{H}_k^{(2)}(b,t)\mathbf{s}_k^{(2)}(b,t) + \mathbf{v}^{(2)}(b,t),$$
$$t \in [KN_t, KN_t + L - 1]. \quad (8)$$

It is assumed that the channel coefficients between two adjacent blocks remain approximately constant (i.e., $\mathbf{H}_k^{(1)}(b-1,t) = \mathbf{H}_k^{(1)}(b,t) = \mathbf{H}_k$). Moreover, when $t \in$ $[(k-1)N_t, kN_t - 1]$, only user k is active, so that the summation operation in Eq. (7) can be removed, i.e.,

$$\mathbf{y}^{(1)}(b,t) = \mathbf{H}_k\mathbf{s}_k^{(1)}(b-1,t)s_k^{(1)}(b,t) + \mathbf{v}^{(1)}(b,t)$$
$$= \mathbf{y}^{(1)}(b-1,t)s_k^{(1)}(b,t) + \tilde{\mathbf{v}}^{(1)}(b,t), \quad (9)$$
$$t \in [(k-1)N_t, kN_t - 1],$$

where $\tilde{\mathbf{v}}^{(1)}(b,t) = \mathbf{v}^{(1)}(b,t) - s_k^{(1)}(b,t)\mathbf{v}^{(1)}(b-1,t)$. Therefore, the transmitted symbol in the first sub-block can be estimated by

$$\hat{s}_k^{(1)}(b,t) = \arg\min_{s_k^{(1)}(b,t)} \left\| \mathbf{y}^{(1)}(b,t) - \mathbf{y}^{(1)}(b-1,t)s_k^{(1)}(b,t) \right\|_F^2,$$
$$t \in [(k-1)N_t, kN_t - 1]. \quad (10)$$

In the second sub-block, $\tilde{t}$ can be estimated by

$$\hat{\tilde{t}} = \arg\max_{i \in [(k-1)N_t, kN_t-1]} \left| \left(\mathbf{y}^{(1)}(b,i)\right)^H \mathbf{y}^{(2)}(b,t) \right|. \quad (11)$$

Then, the active antenna index, $q_k^{(2)}(b,t)$, can be obtained by

$$\hat{q}_k^{(2)}(b,t) = \hat{q}_k^{(1)}(b,\hat{\tilde{t}}) = \text{mod}(\hat{\tilde{t}}, N_t). \quad (12)$$

By obtaining $\hat{\tilde{t}}$, the transmitted M-PSK symbol can be estimated by energy detector as

$$\hat{s}_k^{(1)}(b,t) = \text{demodu}\left(\left(\mathbf{y}^{(1)}\left(b,\hat{\tilde{t}}\right)\right)^H \mathbf{y}^{(2)}(b,t) \right), \quad (13)$$

where $\text{demodu}(\cdot)$ represents the M-PSK demodulation function.

III. THE PROPOSED DIFFERENTIAL STBC-SM SYSTEM

Inspired by the above block-based transmission structure and to further enhance the diversity gain, we propose a novel differential STBC-SM scheme for uplink multi-user massive MIMO communications in the following subsection.

The transmitter block diagram of the proposed differential STBC-SM system for uplink multi-user massive MIMO communications is depicted in Fig. 1. Similar to MU-DSM in [26], the proposed scheme is also performed on a pre-block basis. The specific mapping rule of the proposed transmission structure is explained in the sequel.

A. Mapping rule for the first sub-block

The principle of transmission at the first sub-block is the same as that of the MU-DSM system, which is based on the idea of time division multiple access (TDMA). Specifically, the transmitter employs a pre-designed TA activation pattern, and BS knows this pattern in advance (i.e., the TAs of all users are activated in order during KN_t consecutive time slots). Thus, in the b-th block, the pre-defined antenna pattern matrix of the k-th user, $\mathbf{A}_k^{(1)}(b)$, is a $(N_t \times N_t)$-element identity matrix, i.e., $\mathbf{A}_k^{(1)}(b) = \mathbf{I}_{N_t}$. The antenna matrix for all users, $\mathbf{A}^{(1)}(b)$, can be expressed as

$$\mathbf{A}^{(1)}(b) = \begin{bmatrix} \mathbf{I}_{N_t} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N_t} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I}_{N_t} \end{bmatrix} = \mathbf{I}_{KN_t}. \quad (14)$$
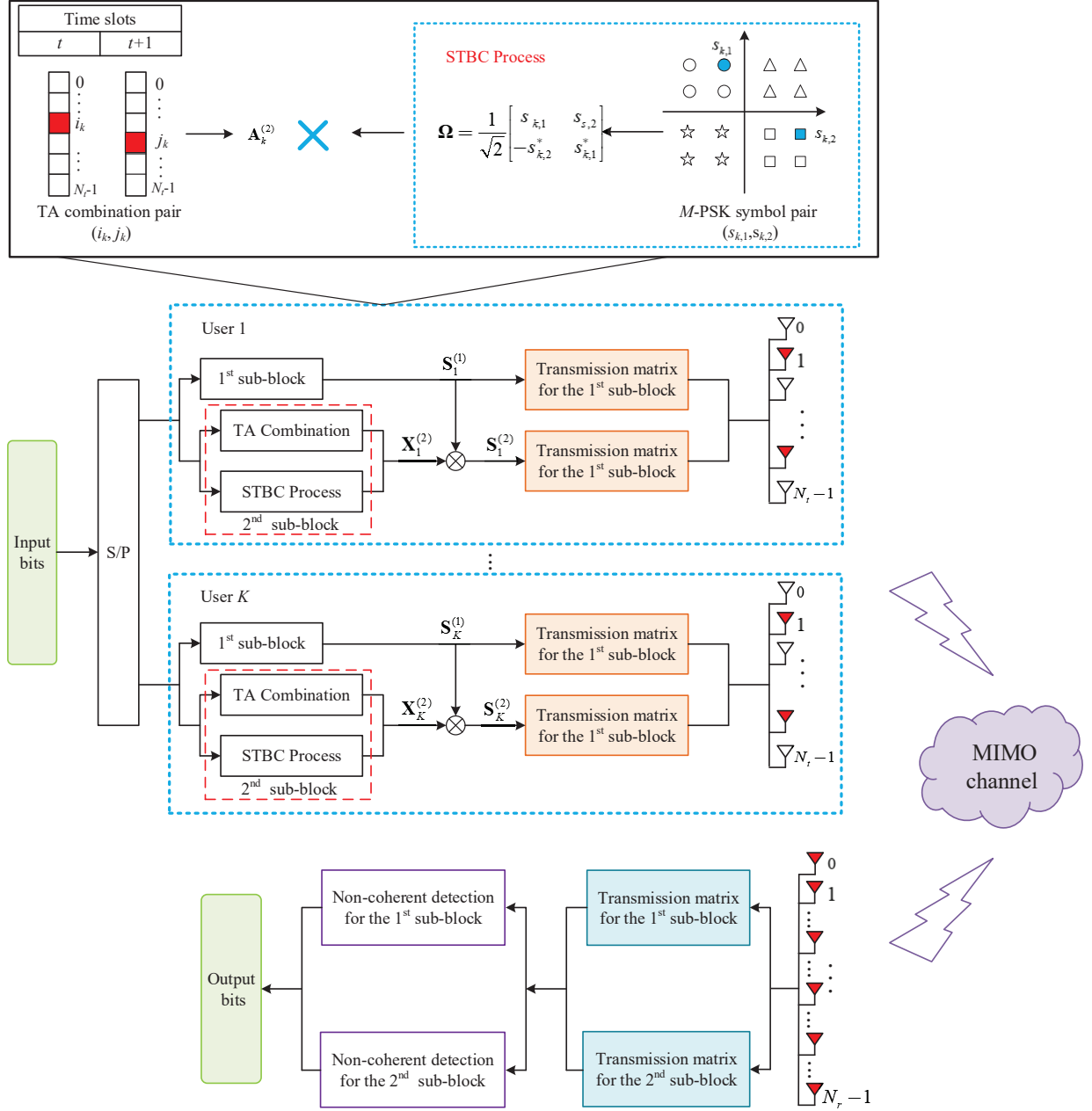
Fig. 1. System model of the proposed differential STBC-SM system for uplink multi-user massive MIMO communications.

In the first sub-block, since the activated antennas obey a pre-defined pattern, which do not transmit bits, all the information bits are conveyed by the M-PSK symbols through KN_t time slots. More specifically, $\mathcal{B}^{(1)} = KN_t\log_2(M)$ bits are modulated into KN_t M-PSK symbols transmitted through the pre-defined TA pattern, which can be written as

$$\mathbf{s}^{(1)}(b) = \left[\mathbf{s}_1^{(1)}(b), ..., \mathbf{s}_k^{(1)}(b), ..., \mathbf{s}_K^{(1)}(b)\right], \quad (15)$$

where

$$\mathbf{s}_k^{(1)}(b) = [s_k^{(1)}(t), s_k^{(1)}(t+1), ..., s_k^{(1)}(t+N_t-1)] \in \mathbb{C}^{1\times N_t}$$

is the modulated M-PSK vector of the k-th user in N_t consecutive time slots. Hence, the transmitted matrix $\mathbf{X}^{(1)}(b)$ before differential coding is

$$\mathbf{X}^{(1)}(b) = \mathbf{A}^{(1)}(b)\text{diag}(\mathbf{s}^{(1)}(b)). \quad (16)$$

Then, differentially encoded transmission matrix of the first sub-block, $\mathbf{S}^{(1)}(b) \in \mathbb{C}^{KN_t \times KN_t}$, can be formed by

$$\mathbf{S}^{(1)}(b) = \mathbf{S}^{(1)}(b-1)\mathbf{X}^{(1)}(b), \quad (17)$$

where $\mathbf{S}^{(1)}(0) = \mathbf{I}_{KN_t}$ is the transmission matrix for the 0-th block. It is worth mentioning that the transmission matrix for each user in the first sub-block, $\mathbf{S}_k^{(1)}(b) \in \mathbb{C}^{N_t \times N_t}$, can be written as

$$\begin{aligned} \mathbf{S}_k^{(1)}(b) &= \mathbf{S}_k^{(1)}(b-1)\mathbf{X}_k^{(1)}(b) \\ &= \mathbf{S}_k^{(1)}(b-1)\mathbf{A}_k^{(1)}(b)diag(\mathbf{s}_k^{(1)}(b)), \end{aligned} \quad (18)$$

which will be used for differential coding in the second sub-block.

Algorithm 1 Mapping Rule of the Second Sub-block for User k of the Proposed Differential STBC-SM Scheme

---

1) First, the information bits $\mathcal{B}_k^{(2)}$ are partitioned into two groups with $\lfloor \log_2 \binom{N_t}{2} \rfloor$ and $2\log_2(M)$ bits, respectively, which are transmitted within one differential STBC-SM block.

2) Next, the first $\lfloor \log_2 \binom{N_t}{2} \rfloor$ bits are used to select a TA combination pair (i_k, j_k) $(i_k, j_k \in [1, 2, ..., N_t]$ and $i_k \neq j_k)$, forming an antenna matrix $\mathbf{A}_k^{(2)}(b, n) \in \mathbb{R}^{N_t \times 2}$ as

$$\mathbf{A}_k^{(2)}(b, n) = [\mathbf{e}_{i_k}, \mathbf{e}_{j_k}], \quad (19)$$

where $n \in [1, 2, ..., L/2]$ and L represent the n-th STBC-SM transmission block and the total time slots in the second sub-block, respectively. Then, the remaining $2\log_2(M)$ bits are modulated to M-PSK symbol pair $(s_{k,1}, s_{k,2})$, which can form the Alamouti's STBC matrix as

$$\mathbf{\Omega} = \frac{1}{\sqrt{2}} \begin{bmatrix} s_{k,1} & s_{k,2} \\ -s_{k,2}^* & s_{k,1}^* \end{bmatrix}. \quad (20)$$

As a result, the STBC-SM matrix is

$$\mathbf{X}_k^{(2)}(b, n) = \mathbf{A}_k^{(2)}(b, n) \mathbf{\Omega} \in \mathbb{C}^{N_t \times 2}. \quad (21)$$

3) Finally, the transmission matrix $\mathbf{S}_k^{(2)}(b, n)$ of user k is obtained via differential coding as

$$\mathbf{S}_k^{(2)}(b, n) = \mathbf{S}_k^{(1)}(b) \mathbf{X}_k^{(2)}(b, n) \in \mathbb{C}^{N_t \times 2}, \quad (22)$$

where $\mathbf{S}_k^{(1)}(b)$ is the transmitted matrix of user k in the first sub-block, defined in Eq. (18).

B. Mapping rule for the second sub-block

As for the second sub-block, we introduce differential STBC-SM transmission scheme to achieve the attainable diversity gain, which is fundamentally different from the system in [26]. Without the loss of generality, we use user k as an example to explain the mapping rule for the second sub-block. Specifically, for user k, in each two consecutive t-th and $(t+1)$-th time slots, information bits $\mathcal{B}_k^{(2)}$ are transmitted by the M-PSK symbol pair $(s_{k,1}, s_{k,2})$ and the activated TA combination pair (i_k, j_k) via every two adjacent time slots. As shown in Fig. 1, the mapping rule for the second sub-block mainly consists of three steps: 1) information bits are first split into two groups; 2) the first group are used to select one of the TA combination pair (i_k, j_k), which can form the antenna matrix $\mathbf{A}_k^{(2)}(b, n)$ to carry STBC symbol. In addition, the remaining group is modulated to M-PSK symbol pair $(s_{k,1}, s_{k,2})$, forming STBC matrix $\mathbf{\Omega}$. Thus, the transmitted STBC-SM matrix can be obtained by combining antenna matrix $\mathbf{A}_k^{(2)}(b, n)$ as well as STBC matrix $\mathbf{\Omega}$; 3) the transmitted STBC-SM matrix $\mathbf{X}_k^{(2)}(b, n)$ will be differentially encoded with $\mathbf{S}_k^{(1)}(b)$ to obtain the final transmitted differential STBC-SM matrix. The procedure of the differential STBC-SM transmission scheme is detailed in Algorithm 1.

Next, we use a simple example to illustrate the detail of mapping rule more intuitively.

Example: Fig. 2 describes an example of the mapping rule for the proposed differential STBC-SM system with $L = 4$, $K = 2$, $N_t = 4$ and QPSK modulation (the mapping rules for antenna combination selection $(N_t = 4)$ and QPSK modulation are shown in Table I). As illustrated in Fig. 2, we can obtain the transmission matrix of the b-th block according to the incoming bit sequence as well as the previous transmission (i.e., the transmission matrix of the $(b-1)$-th block). To be specific, for each user, the total transmit bits of length $\mathcal{B}$ for the b-th block are partitioned into two groups: 1) the first $N_t \log_2(M) = 8$ bits are modulated into the M-PSK symbol transmitted through the pre-designed TA mode in the first sub-block. 2) the second $\frac{KL}{2}(\lfloor \log_2 \binom{N_t}{2} \rfloor + 2\log_2(M)) = 12$ bits are used to select the M-PSK symbol pairs as well as a transmit antenna combination in the second sub-block. More intuitively, for the first user in the b-th block as shown in Fig. 2, the incoming input bits are '11010100000100101110' (total 20 bits for each user in a transmission block). The first 8 bits '11010100' are mapped to the QPSK symbol vector $\mathbf{s}_1^{(1)}(b) = [-j, +j, +j, +1]$, thus, the transmit matrix before differential coding can be expressed as

$$\begin{aligned} \mathbf{X}_1^{(1)}(b) &= \mathbf{A}_1^{(1)}(b) diag(\mathbf{s}_1^{(1)}(b)) \\ &= \begin{bmatrix} -j & 0 & 0 & 0 \\ 0 & +j & 0 & 0 \\ 0 & 0 & +j & 0 \\ 0 & 0 & 0 & +1 \end{bmatrix}. \end{aligned} \quad (23)$$

After differentially encoded with $\mathbf{S}_1^{(1)}(b-1)$, the transmission matrix for the first sub-block can be obtained as

$$\begin{aligned} \mathbf{S}_1^{(1)}(b) &= \mathbf{S}_1^{(1)}(b-1)\mathbf{X}_1^{(1)}(b) \\ &= \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & +j & 0 \\ 0 & 0 & 0 & -j \end{bmatrix} \begin{bmatrix} -j & 0 & 0 & 0 \\ 0 & +j & 0 & 0 \\ 0 & 0 & +j & 0 \\ 0 & 0 & 0 & +1 \end{bmatrix} \\ &= \begin{bmatrix} +j & 0 & 0 & 0 \\ 0 & +j & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -j \end{bmatrix}. \end{aligned} \quad (24)$$

Then, the next 12 bits '000100101110' are used to obtain the transmission matrix for the second sub-block. Since $L = 4$, there are $\frac{L}{2} = 2$ (i.e., $n = 1, 2$) STBC-SM block matrices in the second sub-block. In the first STBC-SM transmission matrix (i.e., $n = 1$), the bit sequence '000100' are divided into two parts, two bits '00' determine the activated antenna combination pair $(0, 1)$, which can form an antenna matrix as

$$\begin{aligned} \mathbf{A}_1^{(2)}(b, 1) &= [\mathbf{e}_0, \mathbf{e}_1] \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned} \quad (25)$$

The four remaining bits '0100' are modulated to the QPSK symbol pair $(+j, +1)$, therefore, the STBC matrix is obtained as

$$\mathbf{\Omega} = \frac{1}{\sqrt{2}} \begin{bmatrix} +j & +1 \\ -1 & -j \end{bmatrix}. \quad (26)$$

	(b-1) TS0	TS1	TS2	TS3	TS4	TS5	TS6	TS7	L TS0 (n=1)	L TS1 (n=1)	L TS2 (n=2)	L TS3 (n=2)	(b) TS0	TS1	TS2	TS3	TS4	TS5	TS6	TS7	L TS0 (n=1)	L TS1 (n=1)	L TS2 (n=2)	L TS3 (n=2)
1st User bit sequence									11 1010		00 0000		11	01	01	00					00 0100		10 1110	
1st User tx – 0th TA	-1								$\frac{-1}{\sqrt2}$	$\frac{-1}{\sqrt2}$	$\frac{-j}{\sqrt2}$	$\frac{-j}{\sqrt2}$	$+j$								$\frac{-1}{\sqrt2}$	$\frac{+j}{\sqrt2}$		
1st TA		$+1$									$\frac{-1}{\sqrt2}$	$\frac{+1}{\sqrt2}$		$+j$							$\frac{-j}{\sqrt2}$	$\frac{+1}{\sqrt2}$	$\frac{-j}{\sqrt2}$	$\frac{+1}{\sqrt2}$
2nd TA			$+j$												-1								$\frac{+j}{\sqrt2}$	$\frac{+1}{\sqrt2}$
3rd TA				$-j$					$\frac{-1}{\sqrt2}$	$\frac{+1}{\sqrt2}$						$-j$								
2nd User tx – 0th TA					$+1$						$\frac{-1}{\sqrt2}$	$\frac{+j}{\sqrt2}$					$+j$				$\frac{+j}{\sqrt2}$	$\frac{-1}{\sqrt2}$		
1st TA						-1			$\frac{-j}{\sqrt2}$	$\frac{-1}{\sqrt2}$								$+j$						
2nd TA							$-j$		$\frac{+j}{\sqrt2}$	$\frac{-1}{\sqrt2}$									-1				$\frac{+j}{\sqrt2}$	$\frac{-1}{\sqrt2}$
3rd TA								$+1$			$\frac{+j}{\sqrt2}$	$\frac{-1}{\sqrt2}$								-1	$\frac{-j}{\sqrt2}$	$\frac{-1}{\sqrt2}$	$\frac{+1}{\sqrt2}$	$\frac{-j}{\sqrt2}$
2nd User bit sequence									10 0100		11 1101		01		10		10		11		11 0001		01 1000	

TA combination / APM symbol

Fig. 2. An example of the mapping rule for $K=2$, $N_t=4$, $L=4$ and QPSK modulation.

Thus, the STBC-SM matrix is

$$\mathbf{X}_1^{(2)}(b,1) = \mathbf{A}_1^{(2)}(b,1)\,\mathbf{\Omega}$$
$$= \frac{1}{\sqrt{2}}\begin{bmatrix} +j & +1 \\ -1 & -j \\ 0 & 0 \\ 0 & 0 \end{bmatrix}. \tag{27}$$

By introducing differential coding, the first differential STBC-SM transmission matrix $\mathbf{S}_1^{(2)}(b,1)$ can be expressed by

$$\mathbf{S}_1^{(2)}(b,1) = \mathbf{S}_1^{(1)}(b)\,\mathbf{X}_1^{(2)}(b,1)$$
$$= \frac{1}{\sqrt{2}}\begin{bmatrix} +j & 0 & 0 & 0 \\ 0 & +j & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -j \end{bmatrix}\begin{bmatrix} +j & +1 \\ -1 & -j \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$
$$= \frac{1}{\sqrt{2}}\begin{bmatrix} -1 & +j \\ -j & +1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}. \tag{28}$$

When $n=2$, the same procedure (i.e., Eqs. (25)-(28)) is carried out. Therefore, we can obtain the second differential STBC-SM block matrix as

$$\mathbf{S}_1^{(2)}(b,2) = \mathbf{S}_1^{(1)}(b)\,\mathbf{X}_1^{(2)}(b,2)$$
$$= \frac{1}{\sqrt{2}}\begin{bmatrix} +j & 0 & 0 & 0 \\ 0 & +j & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -j \end{bmatrix}\begin{bmatrix} 0 & 0 \\ -1 & -j \\ -j & -1 \\ 0 & 0 \end{bmatrix}$$
$$= \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 \\ -j & +1 \\ +j & +1 \\ 0 & 0 \end{bmatrix}. \tag{29}$$

Similarly, the transmission matrix for the second user is obtained. Besides, a novel block-by-block non-coherent detector will be introduced in the next section.

C. Diversity order and coding gain analysis

In this subsection, we analyze the diversity order and coding gain of the proposed scheme with $N_t=4$. For each user, the coding gain can be defined as [34]-[35]

$$G = \min_{\mathbf{X}_k^{(2)}(b,p)\neq\mathbf{X}_k^{(2)}(b,q)} \left|\det(\Delta^H\Delta)\right|, \tag{30}$$

where $\Delta = \mathbf{X}_k^{(2)}(b,p) - \mathbf{X}_k^{(2)}(b,q)$ and $\mathbf{X}_k^{(2)}(b,p)$, $\mathbf{X}_k^{(2)}(b,q) \in \mathbb{C}^{N_t\times 2}$ belong to the transmitted STBC-SM matrices set $\mathcal{X}$. Moreover, when $N_t=4$, the elements in $\mathcal{X}$ are

$$\mathcal{X} = \left\{ \begin{bmatrix} s_{k,1} & s_{k,2} \\ -s_{k,2}^* & s_{k,1}^* \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ s_{k,1} & s_{k,2} \\ -s_{k,2}^* & s_{k,1}^* \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ s_{k,1} & s_{k,2} \\ -s_{k,2}^* & s_{k,1}^* \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} s_{k,1} & s_{k,2} \\ 0 & 0 \\ 0 & 0 \\ -s_{k,2}^* & s_{k,1}^* \end{bmatrix} \right\}. \tag{31}$$

Hence, the value of $\Delta^H\Delta$ can be calculated as

$$\Delta^H\Delta = \begin{cases} \begin{bmatrix} A_1 & B_1 \\ B_1 & C_1 \end{bmatrix}, & \text{if } i_k^p = i_k^q \,\&\, j_k^p = j_k^q \\[6pt] \begin{bmatrix} A_2 & B_2 \\ B_2 & C_2 \end{bmatrix}, & \text{if } i_k^p \neq i_k^q \,\&\, j_k^p \neq j_k^q \\[6pt] \begin{bmatrix} A_3 & B_3 \\ B_3 & C_3 \end{bmatrix}, & \text{if } i_k^p \neq i_k^q \,\&\, j_k^p = j_k^q \\[6pt] \begin{bmatrix} A_4 & B_4 \\ B_4 & C_4 \end{bmatrix}, & \text{if } i_k^p = i_k^q \,\&\, j_k^p \neq j_k^q \end{cases} \tag{32}$$

TABLE I
THE MAPPING RULE FOR ANTENNA COMBINATION SELECTION ($N_t = 4$) AND QPSK MODULATION

Input bits	TA combination	QPSK symbol
00	0 and 1	+1
01	2 and 3	+j
10	1 and 2	-1
11	0 and 3	-j

TABLE II
DIVERSITY ORDER AND CODING GAIN COMPARISONS AMONG THE PROPOSED DIFFERENTIAL STBC-SM, DIFFERENTIAL ALAMOUTI AND CONVENTIONAL DSM SCHEMES.

Schemes (N_t, M)	Proposed scheme (4,4)	Differential Alamouti (2,4)	DSM (4,4)
Rate/bpcu	3.33	3	3.33
Diversity order	2	2	1
Coding gain	16	4	1.64

where

$$A_1 = \left| s_{k,1}^p - s_{k,1}^q \right|^2 + \left| (s_{k,2}^p)^* - (s_{k,2}^p)^* \right|^2,$$
$$B_1 = (s_{k,1}^p - s_{k,1}^q)(s_{k,2}^p - s_{k,2}^q) + (s_{k,2}^p - s_{k,2}^p)^*(s_{k,1}^p - s_{k,1}^q)^*,$$
$$C_1 = \left| s_{k,2}^p - s_{k,2}^q \right|^2 + \left| (s_{k,1}^p)^* - (s_{k,1}^q)^* \right|^2,$$
$$A_2 = \left| s_{k,1}^p \right|^2 + \left| s_{k,1}^q \right|^2 + \left| (s_{k,2}^p)^* \right|^2 + \left| (s_{k,2}^q)^* \right|^2,$$
$$B_2 = s_{k,1}^p s_{k,2}^p + s_{k,1}^q s_{k,2}^q + (s_{k,1}^p s_{k,2}^p)^* + (s_{k,1}^q s_{k,2}^q)^*,$$
$$C_2 = \left| s_{k,2}^p \right|^2 + \left| s_{k,2}^q \right|^2 + \left| (s_{k,1}^p)^* \right|^2 + \left| (s_{k,1}^q)^* \right|^2,$$
$$A_3 = \left| s_{k,1}^p \right|^2 + \left| s_{k,1}^q + ((s_{k,2}^p)^*) \right|^2 + \left| (s_{k,2}^q)^* \right|^2,$$
$$B_3 = s_{k,1}^p s_{k,2}^p - ((s_{k,2}^p)^* - s_{k,1}^q)((s_{k,1}^p)^* - s_{k,2}^q) - (s_{k,1}^q s_{k,2}^p)^*,$$
$$C_3 = \left| s_{k,2}^p \right|^2 + \left| (s_{k,1}^p)^* - s_{k,2}^q \right|^2 + \left| (s_{k,1}^q)^* \right|^2,$$
$$A_4 = \left| s_{k,1}^p - s_{k,1}^q \right|^2 + \left| (s_{k,2}^p)^* \right|^2 + \left| (s_{k,2}^q)^* \right|^2,$$
$$B_4 = (s_{k,1}^p - s_{k,1}^q)(s_{k,2}^p - s_{k,2}^q) - (s_{k,1}^p s_{k,2}^p)^* - (s_{k,1}^q s_{k,2}^q)^*,$$
$$C_4 = \left| s_{k,2}^p - s_{k,2}^q \right|^2 + \left| (s_{k,1}^p)^* \right|^2 + \left| (s_{k,1}^q)^* \right|^2. \tag{33}$$

It can be observed from Eqs. (32) and (33) that the diversity order of our proposed differential STBC-SM scheme is two since the rank of $\Delta^H \Delta$ is two in all cases. As a result, the coding gain can be represented as

$$G = \min_{s_{k,1} \neq s_{k,2}} \left| A_m C_m - B_m^2 \right|, m = 1, 2, 3, 4. \tag{34}$$

To further confirm the benefits of our proposed scheme, Table II compares the diversity order and maximum coding gains among the proposed differential STBC-SM, differential Alamouti and conventional DSM schemes under nearly the same transmission rate. It is clear that our proposed differential STBC-SM scheme can provide the same diversity order as that of the differential Alamouti, and achieve a higher diversity order than the conventional DSM scheme. Futhermore, the proposed scheme achieves much higher coding gain compared to the differential Alamouti and DSM schemes.

IV. NON-COHERENT DETECTOR FOR THE PROPOSED DIFFERENTIAL STBC-SM TRANSMISSION SCHEME

Let $\mathbf{H}_k^{(1)}(b)$ and $\mathbf{H}_k^{(2)}(b,n) \in \mathbb{C}^{N_r \times N_t}$ be the unknown fading channel matrices between user k and the BS in the first and second sub-blocks, respectively. In addition, $\mathbf{V}^{(1)}(b)$ and $\mathbf{V}^{(2)}(b,n) \sim \mathcal{CN}(\mathbf{0}, \sigma_v^2 \mathbf{I}_{N_r})$ are the AWGN matrices. Therefore, the received signals $\mathbf{Y}^{(1)}(b)$ and $\mathbf{Y}^{(2)}(b)$ are presented as

$$\mathbf{Y}^{(1)}(b) = \sqrt{E_s}\mathbf{H}^{(1)}(b)\mathbf{S}^{(1)}(b) + \mathbf{V}^{(1)}(b) \tag{35}$$

and

$$\mathbf{Y}^{(2)}(b,n) = \sqrt{E_s}\sum_{k=1}^{K} \mathbf{H}_k^{(2)}(b,n)\mathbf{S}_k^{(2)}(b,n) + \mathbf{V}^{(2)}(b,n),$$
$$n \in [1, 2, ..., L/2], \tag{36}$$

respectively, where $\sqrt{E_s}$ is the average transmitted symbol energy of each transmission block; and $\mathbf{H}^{(1)}(b) = [\mathbf{H}_1^{(1)}(b), ..., \mathbf{H}_k^{(1)}(b), ..., \mathbf{H}_K^{(1)}(b)]$.

In the first sub-block of our proposed differential STBC-SM system, only M-PSK constellation signals convey information bits while the activated TAs do not transmit bits since they follow a pre-defined pattern. Therefore, only digital symbols (M-PSK symbols) need to be detected in the first sub-block. However, it is necessary to jointly detect spatial symbols and digital symbols in the second sub-block because both activated antenna combinations and M-PSK constellation symbols carry information bits. Next, we will detail the non-coherent detection algorithm for the proposed differential STBC-SM scheme.

A. The first sub-block detection

Since all the transmission symbols are orthogonal in the first sub-block, the transmitted symbols can be estimated by conventional single user non-coherent detector. Based on the assumption that $\mathbf{H}^{(1)}(b-1) = \mathbf{H}^{(1)}(b) = \mathbf{H}$, Eq. (35) is equal to

$$\begin{aligned} \mathbf{Y}^{(1)}(b) &= \sqrt{E_s}\mathbf{H}^{(1)}(b)\mathbf{S}^{(1)}(b) + \mathbf{V}^{(1)}(b) \\ &= \sqrt{E_s}\mathbf{H}\mathbf{S}^{(1)}(b-1)\mathbf{X}^{(1)}(b) + \mathbf{V}^{(1)}(b) \\ &= \sqrt{E_s}\mathbf{Y}^{(1)}(b-1)\mathbf{X}^{(1)}(b) + \tilde{\mathbf{V}}^{(1)}(b), \end{aligned} \tag{37}$$

where $\tilde{\mathbf{V}}^{(1)}(b) = \mathbf{V}^{(1)}(b) - \mathbf{X}^{(1)}(b)\mathbf{V}^{(1)}(b-1)$. Thus, in the first sub-block, the transmitted symbols can be estimated by

$$\hat{\mathbf{X}}^{(1)}(b) = \arg\min_{\mathbf{X}^{(1)}(b)} \left\| \mathbf{Y}^{(1)}(b) - \mathbf{Y}^{(1)}(b-1)\mathbf{X}^{(1)}(b) \right\|_F^2. \tag{38}$$

Since the transmission matrix in the first sub-block is a diagonal matrix as shown in Fig. 2, which indicates there is no inter-symbol and inter-user interferences, the transmitted information bits can be estimated perfectly by utilizing the above detection algorithm. Consequently, errors in the first sub-block are very unlikely. Therefore, we only consider the errors occurred in the second sub-block when analysing BER performance in Section V.

B. The second sub-block detection

The detection for the second sub-block is comprised of two steps: 1) detecting each user's activated antenna combination pair (i_k, j_k), and 2) detecting the M-PSK symbol pair $(s_{k,1}, s_{k,2})$ transmitted from that specific antenna combination.

Step 1: detection of the k-th user's activated antenna combination pair (i_k, j_k) is processed based on the correlation between the received signals of the two sub-blocks. To be specific, Suppose that channel stays constant in a block length, i.e., $\mathbf{H}_k^{(1)}(b) = \mathbf{H}_k^{(2)}(b, n) = \mathbf{H}_k$. Based on the channel asymptotic orthogonality theory in massive MIMO [28], we have

$$\lim_{N_r \to \infty} \frac{1}{N_r} \mathbf{H}_k^H \mathbf{H}_k = \mathbf{I}_{N_t} \qquad (39)$$

and

$$\lim_{\substack{N_r \to \infty, \\ k \neq u}} \frac{1}{N_r} \mathbf{H}_k^H \mathbf{H}_u = \mathbf{0}. \qquad (40)$$

Therefore, the activated antenna combination pair (i_k, j_k) can be estimated by

$$\begin{cases} \hat{i}_k = \mathrm{mod}(\tilde{i}_k, N_t) \\ \hat{j}_k = \mathrm{mod}(\tilde{j}_k, N_t) \end{cases} \qquad (41)$$

where

$$\begin{aligned} &(\tilde{i}_k, \tilde{j}_k) \\ &= \underset{\substack{i_k, j_k \in \{[(k-1)N_t, kN_t - 1]\} \\ i_k \neq j_k}}{\arg\max} \left| (\mathbf{Y}^{(1)}(b)\{i_k, j_k\})^H \mathbf{Y}^{(2)}(b, n) \right|, \\ &k = 1, 2, ..., K. \end{aligned} \qquad (42)$$

and $\mathbf{Y}^{(1)}(b)\{i_k, j_k\}$ refers to the i_k-th and j_k-th columns of $\mathbf{Y}^{(1)}(b)$.

It is evident that the power of the signals received from the active TA combinations is significantly higher than that of the inactive TAs. Therefore, we can distinguish indices of the active antennas from that of the inactive antennas according to the power differences of the right side of Eq. (42) (i.e. $(\mathbf{Y}^{(1)}(b)\{i_k, j_k\})^H \mathbf{Y}^{(2)}(b, n)$). More specifically, the greater the absolute value of $(\mathbf{Y}^{(1)}(b)\{i_k, j_k\})^H \mathbf{Y}^{(2)}(b, n)$, the stronger the signal power, which means that the corresponding antenna is more likely to be activated.

Step 2: The same as the detection of conventional STBC-SM, the transmitted M-PSK symbol pair $(s_{k,1}, s_{k,2})$ is able to be estimated by obtaining the estimated antenna combination pair $(\hat{i}_k, \hat{j}_k)$. Thus, the transmitted M-PSK symbol pair $(s_{k,1}, s_{k,2})$ can be estimated as

$$(\hat{s}_{k,1}, \hat{s}_{k,2}) = \underset{s_{k,1}, s_{k,2}}{\arg\min} \left\| \mathbf{Y}^{(2)}(b, n) - \mathbf{Y}_k^{(1)}(b) \mathbf{A}_k^{(2)}(b, n) \mathbf{\Omega} \right\|_F^2, \qquad (43)$$

which can be further simplified as

$$(\hat{s}_{k,1}, \hat{s}_{k,2}) = \underset{s_{k,1}, s_{k,2}}{\arg\min} \left\| \mathbf{Y}^{(2)}(b, n) - (\mathbf{Y}^{(1)}(b)\{\tilde{i}_k, \tilde{j}_k\}) \mathbf{\Omega} \right\|_F^2, \qquad (44)$$

where $\mathbf{Y}_k^{(1)}(b) = \mathbf{H}_k^{(1)}(b) \mathbf{S}_k^{(1)}(b) + \mathbf{V}_k^{(1)}(b)$ is the received signal for user k in the first sub-block of the b-th block.

Observed from the above detection algorithm that the proposed detection approach is independent at different time slots, errors are not propagated to other time slots which is in contrast to the conventional STBC-DSM/STBC-RDSM scheme.

C. Complexity analysis

In this subsection, we provide analysis on the computational complexity of the proposed block-by-block non-coherent detection in terms of complex multiplications.

The complexity of the detection for the first sub-block is mainly due to Eq. (38), which can be written as

$$C^{(1)} = (KN_t + KN_rN_t) \times M, \qquad (45)$$

where KN_t is the computation required for calculating $\mathbf{Y}^{(1)}(b - 1)\mathbf{X}^{(1)}(b)$ and KN_rN_t is the number of operations required for calculating the Frobenius norm $\left\| \mathbf{Y}^{(1)}(b) - \mathbf{Y}^{(1)}(b-1)\mathbf{X}^{(1)}(b) \right\|_F^2$. The computation of the above terms is repeated M times.

As for the second sub-block, the complexity can be calculated by

$$C^{(2)} = KL(2N_r + 2) + (3KLN_r) \times M^2, \qquad (46)$$

where $KL(2N_r + 2)$ is for estimating the activated antenna combination i.e, Eq. (42). Furthermore, the number of operations required for estimating the transmitted M-PSK symbol pair in Eq. (44) is $(3KLN_r)M^2$.

Therefore, the total computational complexity of the proposed non-coherent detection is

$$\begin{aligned} C &= C^{(1)} + C^{(2)} \\ &= (KN_t + KN_rN_t)M + KL(2N_r + 2) + (3KLN_r)M^2 \end{aligned} \qquad (47)$$

Table III illustrates the computational complexity comparison between the proposed non-coherent detector and its coherent counterpart (i.e., coherent detector with the least squares (LS) channel estimation) under various parameter settings. It can be seen from Table III that although the complexity of the proposed non-coherent detector is slightly higher than that of the LS channel estimation based coherent detector, more information bits are transmitted in the proposed non-coherent scheme. To be specific, in the case of $K = 4$, $N_t = 4$, $L = 4$ and QPSK, the proposed non-coherent scheme transmits 32 more information bits than the coherent scheme, resulting in enhanced spectrum efficiency.

V. PERFORMANCE ANALYSIS

This section first compares the transmission rate between the proposed differential STBC-SM scheme and conventional MU-DSM in [26]. Subsequently, an average bit error probability (ABEP) upper bound is derived based on the non-coherent detection.

A. Transmission rate comparison

The transmission rate for each block of the proposed differential STBC-SM system is

$$R_{\mathrm{Proposed}} = KN_t \log_2(M) + \frac{KL}{2} (\lfloor \log_2 \binom{N_t}{2} \rfloor + 2\log_2(M)), \qquad (48)$$

Schemes (N_t, N_r, K, L, M)	Proposed non-coherent detector	coherent detector with LS channel estimation
(4,32,2,4,2)	4128	3856
(4,64,4,4,4)	55392	52256
(4,32,8,4,8)	207168	199744

TABLE IV
THE TRANSMISSION RATE OF THE PROPOSED DIFFERENTIAL STBC-SM
AND CONVENTIONAL DSM IN [26] SCHEMES WITH DIFFERENT
CONFIGURATIONS

Schemes	$K = 4$, $N_t = 4$, $L = 4$, QPSK	$K = 2$, $N_t = 4$, $L = 4$, BPSK
Proposed differential STBC-SM	80 bits/block	24 bits/block
Conventional DSM in [26]	96 bits/block	32 bits/block

where $KN_t \log_2(M)$ is the total transmit bits for the first sub-block because the TAs obey a pre-designed active pattern. $\frac{KL}{2}(\lfloor \log_2 \binom{N_t}{2} \rfloor + 2\log_2(M))$ denotes all the information bits carried by M-PSK symbols as well as spatial antenna combination in the second sub-block.

Moreover, the block transmission rate of the conventional MU-DSM in [26] is calculated by

$$R_{\text{DSM}} = KN_t \log_2(M) + KL(\log_2(N_t M)). \quad (49)$$

Observed from Eqs. (48)-(49), the transmission rate in the first sub-block of the proposed transmission structure is the same as that of the conventional MU-DSM. However, in the second sub-block, the proposed system loses part of the transmission rate compared to the conventional MU-DSM scheme, since it utilizes STBC structure to obtain higher diversity gain.

Table IV depicts the transmission rate of these two schemes with different parameter settings. More intuitively, the proposed scheme transfers 16 bits less than the traditional DSM scheme in the case of $K = 4$, $N_t = 4$, $L = 4$ and QPSK modulation, but it achieves considerable BER performance improvement when the transmission rate is the same, which will be evidence by our simulation results.

B. BER performance analysis

The average BER of the proposed differential STBC-SM scheme is given by

$$P_e = \frac{1}{K} \sum_{k=1}^{K} P_{e,k}, \quad (50)$$

where $P_{e,k}$ denotes the overall BER of user k. Next, we will focus on the derivation of the BER performance of user k.

Assuming that each user transfers m bits during two adjacent block intervals. The differential STBC-SM data transmission matrices for user k are denoted by $\mathbf{S}_1^{(2)}, \mathbf{S}_2^{(2)}, ..., \mathbf{S}_{2^m}^{(2)}$. According to the well-known union bound technique [29], $P_{e,k}$ in Eq. (50) can be expressed as

$$P_{e,k} \le \frac{1}{m 2^m} \sum_{i=1}^{2^m} \sum_{\substack{j=1, \\ j \ne i}}^{2^m} d(\mathbf{S}_i^{(2)}, \mathbf{S}_j^{(2)}) P(\mathbf{S}_i^{(2)} \to \mathbf{S}_j^{(2)}), \quad (51)$$

where $d(\mathbf{S}_i^{(2)}, \mathbf{S}_j^{(2)})$ refers to the Hamming distance between $\mathbf{S}_i^{(2)}$ and $\mathbf{S}_j^{(2)}$, and $P(\mathbf{S}_i^{(2)} \to \mathbf{S}_j^{(2)})$ is the PEP.

In order to simplify the expression, we use $\mathbf{Y}_k^{(1)}$, $\mathbf{Y}^{(2)}$, $\mathbf{V}_k^{(1)}$ and $\tilde{\mathbf{V}}^{(2)}$ replace $\mathbf{Y}_k^{(1)}(b)$, $\mathbf{Y}^{(2)}(b, n)$, $\mathbf{V}_k^{(1)}(b)$ and $\tilde{\mathbf{V}}^{(2)}(b, n)$ in the remainder of this paper, respectively. Based on the non-coherent detection algorithm in Eq. (44), the error will occur when

$$\left\| \mathbf{Y}^{(2)} - \mathbf{Y}_k^{(1)} \mathbf{X}_j^{(2)} \right\|_F^2 \le \left\| \mathbf{Y}^{(2)} - \mathbf{Y}_k^{(1)} \mathbf{X}_i^{(2)} \right\|_F^2, \quad (52)$$

which is equal to

$$\mathbf{Y}_k^{(1)} \mathbf{D}_{i,j} (\mathbf{Y}_k^{(1)})^H \le 2\text{Re}\left\{ (\tilde{\mathbf{V}}^{(2)})^H \mathbf{\Delta}_{i,j} \mathbf{Y}_k^{(1)} \right\}, \quad (53)$$

with $\mathbf{\Delta}_{i,j} = \mathbf{I} \otimes (\mathbf{X}_j^{(2)} - \mathbf{X}_i^{(2)})$ and $\mathbf{D}_{i,j} = \mathbf{\Delta}_{i,j} \mathbf{\Delta}_{i,j}^H = \mathbf{I} \otimes ((\mathbf{X}_i^{(2)} - \mathbf{X}_j^{(2)})(\mathbf{X}_i^{(2)} - \mathbf{X}_j^{(2)})^H)$ is the code distance matrix. The derivation from Eqs. (52) to (53) is given in Appendix A.

Let $\Psi = 2\text{Re}\{(\tilde{\mathbf{V}}^{(2)})^H \mathbf{\Delta}_{i,j} \mathbf{Y}_k^{(1)}\}$, which is a Gaussian random variable. For the given $\mathbf{X}_i^{(2)}$ in k-th block, the conditional mean and variance of Ψ are

$$\bar{m} = E\left\{ \Psi \left| \mathbf{Y}_k^{(1)} \right. \right\}$$
$$= 2\text{Re}\left\{ (\mathbf{Y}_k^{(1)})^H \mathbf{\Delta}_{i,j}^H \mathbf{X}_i^{(2)} \bar{m}_{\mathbf{V}_k^{(1)} | \mathbf{Y}_k^{(1)}} \right\}, \quad (54)$$

and

$$\sigma^2 = E\left\{ \|\Psi - \bar{m}\|^2 \left| \mathbf{Y}_k^{(1)} \right. \right\}$$
$$= 2(\mathbf{Y}_k^{(1)})^H \mathbf{\Delta}_{i,j}^H (\mathbf{I} + \mathbf{X}_i^{(2)} \mathbf{C}_{\mathbf{V}_k^{(1)} | \mathbf{Y}_k^{(1)}} (\mathbf{X}_i^{(2)})^H) \mathbf{\Delta}_{i,j} \mathbf{Y}_k^{(1)}. \quad (55)$$

where $\bar{m}_{\mathbf{V}_k^{(1)} | \mathbf{Y}_k^{(1)}}$ and $\mathbf{C}_{\mathbf{V}_k^{(1)} | \mathbf{Y}_k^{(1)}}$ are the conditional mean and variance of $\mathbf{V}_k^{(1)}$, respectively. By applying the result available in [30], we obtain

$$\bar{m}_{\mathbf{V}_k^{(1)} | \mathbf{Y}_k^{(1)}} = (\frac{E_s}{\sigma_v^2} \mathbf{X}^{(1)} \mathbf{R}(\mathbf{X}^{(1)})^H + \mathbf{I})^{-1} \mathbf{Y}_k^{(1)}, \quad (56)$$

and

$$\mathbf{C}_{\mathbf{V}_k^{(1)} | \mathbf{Y}_k^{(1)}} = \sigma_v^2 \mathbf{I} - \sigma_v^2 (\frac{E_s}{\sigma_v^2} \mathbf{X}^{(1)} \mathbf{R}(\mathbf{X}^{(1)})^H + \mathbf{I})^{-1}, \quad (57)$$

where $\mathbf{R} = E\{\mathbf{H}^H \mathbf{H}\}$ is the correlation matrix of $\mathbf{H}$.

Defining $d_{i,j}^2 = \mathbf{Y}_k^{(1)} \mathbf{D}_{i,j} (\mathbf{Y}_k^{(1)})^H$, the conditional PEP is then presented by

$$P(\mathbf{X}_i^{(2)} \to \mathbf{X}_j^{(2)} | \mathbf{Y}_k^{(1)}) = P(\Psi \ge d_{i,j}^2)$$
$$= \int_{d_{i,j}^2}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\Psi - \bar{m})^2}{2\sigma^2} \right) d\Psi$$
$$= Q\left(\frac{d_{i,j}^2 - \bar{m}}{\sigma} \right), \quad (58)$$

where $Q(x) = \frac{1}{\pi} \int_0^{\pi/2} \exp(-x^2 / 2\sin^2\theta) d\theta$ represents the Gaussian Q-function.

Consequently, the PEP can be further formulated as

$$P(\mathbf{X}_i^{(2)} \to \mathbf{X}_j^{(2)}) = \frac{1}{\pi} \int_0^{\pi/2} \int_{-\infty}^{+\infty} \exp\left(-\frac{\Gamma}{2\sin^2\theta}\right) P_\Gamma(\Gamma) d\Gamma d\theta$$

$$= \frac{1}{\pi} \int_0^{\pi/2} M_\Gamma\left(-\frac{1}{2\sin^2\theta}\right) d\theta \tag{59}$$

where $\Gamma = \frac{(d_{i,j}^2 - \bar{m})^2}{\sigma^2}$, $M_\Gamma(s) = \int_0^{\pi/2} P_\Gamma(\Gamma) \exp(s\Gamma) d\Gamma$ denotes the MGF of random variable Γ.

Based on the channel asymptotic orthogonality theory in massive MIMO as shown in Eq. (39), we have $\mathbf{R} = \mathbf{I}$. Hence, the mean and variance of $\mathbf{V}_k^{(1)}$ conditioned on $\mathbf{Y}_k^{(1)}$ can be simplified to

$$\bar{m}_{\mathbf{V}_k^{(1)}|\mathbf{Y}_k^{(1)}} = \frac{\sigma_v^2}{(\sigma_v^2 + E_s)}\mathbf{Y}_k^{(1)} \tag{60}$$

and

$$\mathbf{C}_{\mathbf{V}_k^{(1)}|\mathbf{Y}_k^{(1)}} = \frac{E_s\sigma_v^2}{(\sigma_v^2 + E_s)}\mathbf{I}. \tag{61}$$

Consequently, $\bar{m}$ and σ^2 can be rewritten as

$$\begin{aligned}\bar{m} &= \frac{2\sigma_v^2}{(\sigma_v^2 + E_s)} \mathrm{Re}\left\{(\mathbf{Y}^{(1)})^H \boldsymbol{\Delta}_{i,j}^H \mathbf{X}_i^{(2)} \mathbf{Y}^{(1)}\right\} \\ &= \frac{2\sigma_v^2}{(\sigma_v^2 + E_s)} \mathrm{Re}\left\{(\mathbf{Y}_k^{(1)})^H (\mathbf{I} - (\mathbf{X}_j^{(2)})^H \mathbf{X}_i^{(2)}) \mathbf{Y}_k^{(1)}\right\} \\ &= \frac{\sigma_v^2}{(\sigma_v^2 + E_s)} (\mathbf{Y}_k^{(1)})^H \mathbf{D}_{i,j} \mathbf{Y}_k^{(1)} \\ &= \frac{\sigma_v^2}{(\sigma_v^2 + E_s)} d_{i,j}^2\end{aligned} \tag{62}$$

and

$$\begin{aligned}\sigma^2 &= \frac{2(\sigma_v^2 + E_s + E_s\sigma_v^2)}{\sigma_v^2 + E_s}(\mathbf{Y}_k^{(1)})^H \boldsymbol{\Delta}_{i,j}^H \boldsymbol{\Delta}_{i,j} \mathbf{Y}_k^{(1)} \\ &= \frac{2(\sigma_v^2 + E_s + E_s\sigma_v^2)}{\sigma_v^2 + E_s} d_{i,j}^2.\end{aligned} \tag{63}$$

By substituting $\bar{m}$ and σ^2 into Γ, we have

$$\begin{aligned}\Gamma &= \frac{(d_{i,j}^2 - \bar{m})^2}{\sigma^2} \\ &= \frac{E_s^2}{2(\sigma_v^2 + E_s)(\sigma_v^2 + E_s + E_s\sigma_v^2)} d_{i,j}^2.\end{aligned} \tag{64}$$

Since $d_{i,j}^2$ is a quadratic form, whose MGF has already been derived in [31], we can obtain the MGF of Γ as

$$M_\Gamma(s) = \left|\mathbf{I} - \frac{sE_s^2 \Phi(\mathbf{D}_{i,j})^H \mathbf{D}_{i,j}}{2(\sigma_v^2 + E_s)(\sigma_v^2 + E_s + E_s\sigma_v^2)}\right|^{-1}, \tag{65}$$

where $\Phi = E\{(\mathbf{Y}_k^{(1)})^H \mathbf{Y}_k^{(1)}\}$ is the covariance matrix of the received signal in the first sub-block. For Rayleigh flat fading channel, the covariance matrix is $\Phi = (E_s + \sigma_v^2)\mathbf{I}$. Thus, the MGF of Γ can be rewritten as

$$M_\Gamma(s) = \prod_{l=1}^L \left(1 - \frac{sE_s^2(\sigma_v^2 + E_s)\lambda_l}{2(\sigma_v^2 + E_s)(\sigma_v^2 + E_s + E_s\sigma_v^2)}\right)^{-N_r}, \tag{66}$$

where L is the number of the distinct non-zero eigenvalues λ_l of the distance matrix $(\mathbf{D}_{i,j})^H \mathbf{D}_{i,j}$ with multiplicity N_r.

Therefore, the PEP can be derived as follows

$$P(\mathbf{X}_i^{(2)} \to \mathbf{X}_j^{(2)}) = \frac{1}{\pi} \int_0^{\pi/2} \prod_{l=1}^L \left(\frac{1}{1 + \frac{\beta\lambda_l}{2\sin^2\theta}}\right)^{N_r} d\theta, \tag{67}$$

where $\beta = \frac{E_s^2(\sigma_v^2 + E_s)}{(\sigma_v^2 + E_s)(\sigma_v^2 + E_s + E_s\sigma_v^2)}$. Taking advantage of the expansion equation in [32, Eq. (40)], the product in Eq. (67) can be rewritten as

$$\prod_{l=1}^L \left(\frac{1}{1 + \frac{\beta\lambda_l}{2\sin^2\theta}}\right)^{N_r} = \sum_{l=1}^L \sum_{p=1}^{N_r} A_{l,p} \left(\frac{1}{1 + \frac{\beta\lambda_l}{2\sin^2\theta}}\right)^p, \tag{68}$$

where

$$A_{l,p} = \frac{\left\{\frac{p^{N_r - p}}{px^{N_r - p}} \prod_{n=1, n\neq l}^L \left(\frac{1}{1 + x\lambda_n}\right)^p\right\}|x = -\lambda_l^{-1}}{(N_r - p)!(\lambda_l)^{N_r - p}}. \tag{69}$$

Based on the integral formula in [29]

$$\begin{aligned}&\frac{1}{\pi} \int_0^{\pi/2} \left(\frac{\sin^2\theta}{\sin^2\theta + c}\right)^m d\theta \\ &= \frac{1}{2}\left[1 - \omega(c) \sum_{k=0}^{m-1} \binom{2k}{k}\left(\frac{1 - \omega^2(c)}{4}\right)^k\right],\end{aligned} \tag{70}$$

where $\omega(c) = \sqrt{\frac{c}{1+c}}$, we can further derive the PEP in a closed form as

$$\begin{aligned}P(\mathbf{X}_i^{(2)} \to \mathbf{X}_j^{(2)}) &= \frac{1}{\pi} \int_0^{\pi/2} \sum_{l=1}^L \sum_{p=1}^{N_r} A_{l,p} \left(\frac{1}{1 + \frac{\beta\lambda_l}{2\sin^2\theta}}\right)^p d\theta \\ &= \sum_{l=1}^L \sum_{p=1}^{N_r} A_{l,p} \left\{\frac{1}{\pi} \int_0^{\pi/2} \left(\frac{\sin^2\theta}{\sin^2\theta + \beta\lambda_l/2}\right)^p d\theta\right\} \\ &= \frac{1}{2}\sum_{l=1}^L \sum_{p=1}^{N_r} A_{l,p} \left[1 - \sqrt{\frac{\beta\lambda_l}{2 + \beta\lambda_l}} \sum_{j=1}^p \binom{2j}{j}\left(\frac{1}{4 + 2\beta\lambda_l}\right)^j\right].\end{aligned} \tag{71}$$

C. A practical scenario

Note that the time slots required in the first sub-block of the proposed differential STBC-SM scheme is KN_t, which is the same as the minimum length of the orthogonal pilots in conventional coherent uplink systems [36]. In the proposed scheme, a portion of information bits can be transmitted during the first sub-block, while conventional coherent schemes transmit known symbols (i.e., pilots) for channel estimation. In other words, the proposed differential STBC-SM scheme can provide higher transmission rate than its coherent counterpart. As the value of K and N_t increases, this benefit becomes more obvious.

In addition, in practical scenarios, the length of the first sub-block is negligible compared to the total transmission block length. It is evident that the Maximum Doppler spread f_m and

coherence time T_c are inversely proportional to one another, i.e.,

$$T_c \approx \frac{1}{f_m}. \tag{72}$$

Since the maximum Doppler frequency in LTE channel model is 5 Hz, 70 Hz and 300 Hz in low, medium and high mobility scenarios [37, Table 11.2], respectively. As a result, under these LTE systems, the coherence time T_c are 200 ms, 14 ms and 3 ms, respectively. In high mobility scenarios (i.e., T_c = 3 ms), supposed that the maximum bit rate per user is 10 Mbps (i.e., the symbol duration T is 10 ns), the total time slots contained in the channel's coherence time can be calculated by

$$N = \frac{T_c}{T} = 300000. \tag{73}$$

Since the total length of a transmission block must be a small fraction of N, we assume that the transmission block length of our proposed scheme is one-tenth of N, i.e., $KN_t + L = \frac{N}{10}$, the time slots contained in our proposed scheme can thus be calculated as

$$KN_t + L = 30000. \tag{74}$$

To make this issue more intuitively, let us consider uplink transmissions with $K = 100$ and $N_t = 4$, the time slots required in the first sub-block is 400, which only accounts for 1% of the total block length. Therefore, in practical scenarios, the length of the first sub-block is negligible compared to the total block length.

VI. NUMERICAL RESULTS AND PERFORMANCE COMPARISON

In this section, we evaluate the BER performance of the proposed differential STBC-SM scheme by Monte Carlo simulations. In all the simulations, the modulation is QPSK and the number of simulations is 10^6 at each N_r or SNR.

In Fig. 3, the theoretical and simulation results of the proposed differential STBC-SM system with different SNR values are illustrated. Fig. 3 shows that the BER upper bound, derived in Section V, becomes tight as the SNR and N_r values increase for the proposed differential STBC-SM system. This validates the accuracy of our theoretical analysis which can be employed as an effective tool to estimate the BER performance behaviour of our proposed differential STBC-SM system.

Fig. 4 depicts the impact of the number of users on the BER performance. The simulation parameters are $N_t = 4$, $L = 4$, and SNR = 15 dB. It demonstrates that the BER performance will degrade with increasing number of users. Obviously, the more users are simultaneously transmitting, the higher the inter-user interference they experience. Consequently, as the number of users increases, the number of RAs should increase accordingly so as to maintain the BER performance. As an example, for a target BER = 10^{-6}, the number of RAs required at the BS is 220, 450 and 850 when $K = 8$, 16 and 32, respectively.

Fig. 5 compares the BER performance among the differential STBC-SM, conventional differential Alamouti in [33] and MU-DSM in [26] under nearly the same transmission rate. In the case of $K = 2$ and $L = 4$, we set the simulation
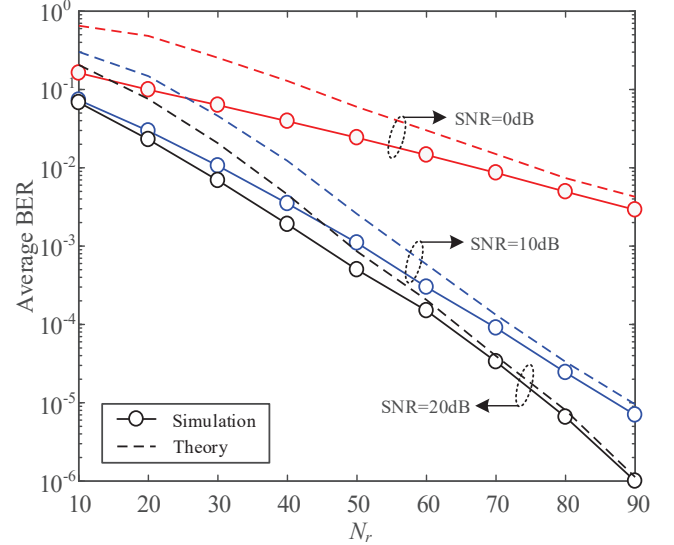

Fig. 3. Simulation and theoretical results of the proposed differential STBC-SM system with $K = 4$, $N_t = 4$, $L = 4$ and QPSK modulation.
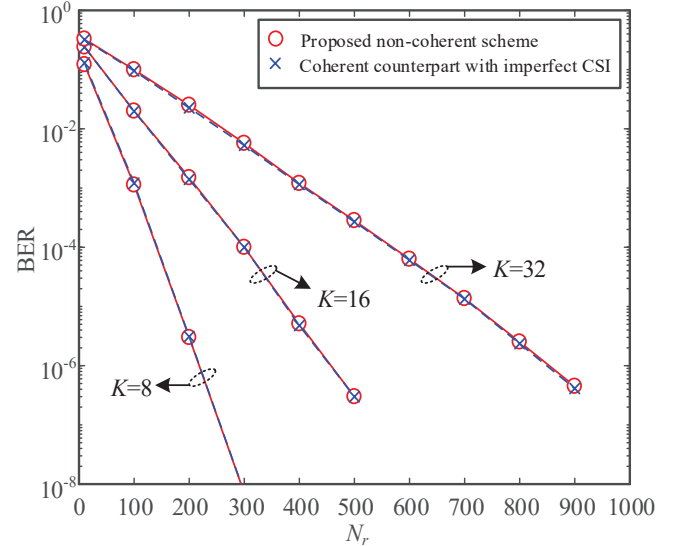

Fig. 4. BER performance of the proposed differential STBC-SM scheme with various numbers of user for the simulation parameters of $N_t = 4$, $L = 4$, SNR = 15 dB and QPSK modulation.

parameters as $N_t = 4$, $N_r = 100$ and $M_1 = M_2 = 4$ in the proposed differential STBC-SM scheme to achieve $40/12 \approx 3.33$ bpcu, $N_t = 4$, $N_r = 100$ and $M_1 = 2$, $M_2 = 4$ in the conventional MU-DSM to obtain $40/12 \approx 3.33$ bpcu and $N_t = 2$, $N_r = 100$ and $M_1 = M_2 = 4$ in differential Alamouti scheme to achieve $24/8 = 3$ bpcu. It is shown in Fig. 5 that the proposed differential STBC-SM scheme provides nearly 2 dB and 4 dB performance gains at BER = 10^{-6} over the conventional differential Alamouti and MU-DSM schemes, respectively. Moreover, when the transmission rate is 4 bpcu, the same tendency can be observed, this is because our proposed differential STBC-SM scheme combines differential coding and STBC-SM transmission structure, which can exploit the diversity gain inherent in STBC while maintaining the benefits of DSM.

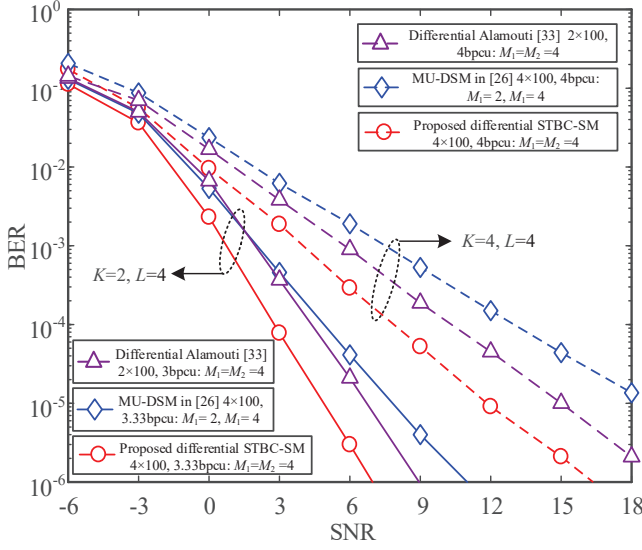Fig. 6 compares the BER performance between the proposed

Fig. 5. BER performance comparisons among the proposed differential STBC-SM, conventional MU-DSM [26] and differential Alamouti [33] under various simulation parameters.
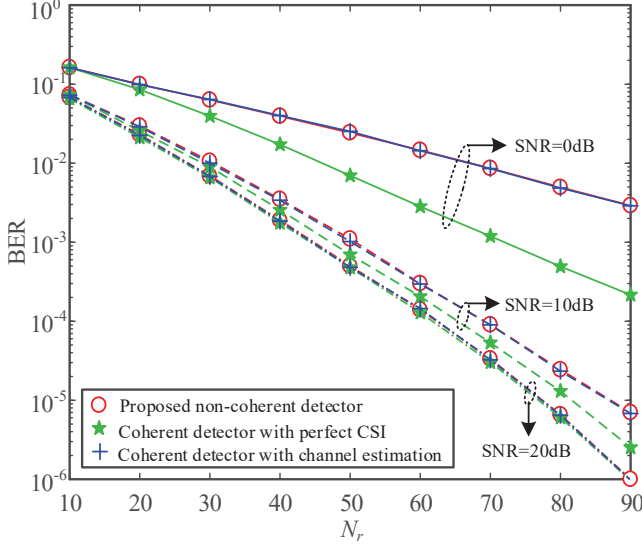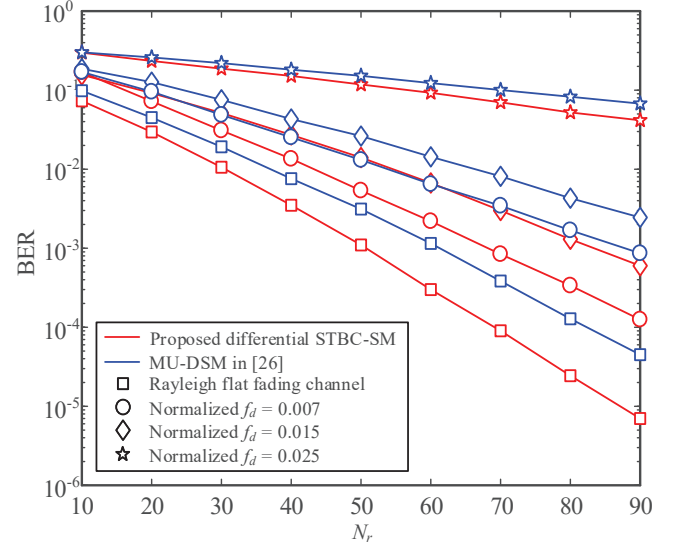

Fig. 7. BER performance comparison between the proposed differential STBC-SM and MU-DSM in [26] under different channel model (Rayleigh fading channel and time-varying channel with different doppler shift f_d) for the case of $K = 4$, $N_t = 4$, $L = 4$, SNR = 10 dB.

channel and the time-varying channels with a different doppler shift f_d) for the parameters of $L = 4$, $K = 4$, $N_t = 4$, and SNR = 10 dB. It can be seen from Fig. 7 that the proposed differential STBC-SM scheme can still provide considerable BER performance improvement under time-varying channels compared to the conventional DSM in [26]. Furthermore, the proposed differential STBC-SM scheme under the high speed channels exhibit worse BER performance than that under the Rayleigh fading channel due to the doppler effect. The performance further deteriorates as the doppler shift f_d increases. However, it is worth mentioning that the coherence time becomes shorter when f_d increases, resulting in excessive pilot overhead when coherent schemes are applied. This justifies the use of our proposed differential STBC-SM algorithm in high mobility scenarios.

Fig. 6. BER performance comparisons between the proposed non-coherent detector and its coherent counterpart (coherent detection with imperfect and perfect CSI) when $K = 4$, $N_t = 4$, $L = 4$ and QPSK modulation.

non-coherent detector and its coherent counterpart (coherent detector with imperfect and perfect CSI) under various parameter configurations. As can be observed from Fig. 6, the proposed non-coherent detector achieves the identical BER performance as the coherent detector with channel estimation for the whole SNR and N_r regions. Moreover, in the high SNR region (i.e., SNR = 20 dB), the BER performance of the proposed non-coherent detector is infinitely close to that of the coherent detection with perfect CSI. It is worth mentioning that our proposed non-coherent transmission scheme can transmit $KN_t\log_2(M)$ additional bits compared to its coherent counterpart, since the M-PSK symbols are transmitted in the first sub-block in the proposed non-coherent scheme.

Fig. 7 compares the BER performance between the proposed differential STBC-SM and the conventional MU-DSM in [26] under different channels (i.e., the static Rayleigh fading

VII. CONCLUSION

A novel differential STBC-SM transmission system is proposed for uplink multi-user massive MIMO communications, which combines the differential coding and STBC-SM to reap their respective benefits. Moreover, a non-coherent detector was presented, which can eliminate the effect of multiuser interference while avoiding error propagation. Furthermore, the derivation of the BER performance based on non-coherent detector is presented. Our simulation results demonstrated that our proposed differential STBC-SM scheme is able to acquire significantly BER performance improvement with acceptable transmission rate loss compared to the conventional DSM scheme without the knowledge of CSI.

APPENDIX A
PROOF OF EQ. (53)

By expanding Eq. (52), we obtain

$$(\mathbf{Y}^{(2)} - \mathbf{Y}_k^{(1)}\mathbf{X}_j^{(2)})^H(\mathbf{Y}^{(2)} - \mathbf{Y}_k^{(1)}\mathbf{X}_j^{(2)}) \leq (\mathbf{Y}^{(2)} - \mathbf{Y}_k^{(1)}\mathbf{X}_i^{(2)})^H(\mathbf{Y}^{(2)} - \mathbf{Y}_k^{(1)}\mathbf{X}_i^{(2)}), \quad (75)$$

$$\Rightarrow (\mathbf{Y}^{(2)})^H \mathbf{Y}_k^{(1)} \mathbf{X}_j^{(2)} + (\mathbf{X}_j^{(2)})^H (\mathbf{Y}_k^{(1)})^H \mathbf{Y}^{(2)} \geq (\mathbf{Y}^{(2)})^H \mathbf{Y}_k^{(1)} \mathbf{X}_i^{(2)} + (\mathbf{X}_i^{(2)})^H (\mathbf{Y}_k^{(1)})^H \mathbf{Y}^{(2)}. \tag{76}$$

Furthermore, we have

$$\begin{aligned} \mathbf{Y}^{(2)}(b,n) &= \sum_{k=1}^{K} \mathbf{H}_k^{(2)}(b,n) \mathbf{S}_k^{(2)}(b,n) + \mathbf{V}^{(2)}(b,n) \\ &= \sum_{k=1}^{K} \mathbf{H}_k^{(2)} \mathbf{S}_k^{(1)}(b) \mathbf{X}_k^{(2)}(b,n) + \mathbf{V}^{(2)}(b,n) \\ &= \mathbf{Y}_k^{(1)}(b) \mathbf{X}_k^{(2)}(b,n) + \sum_{i=1,i\neq j}^{K} \mathbf{H}_i^{(2)} \mathbf{S}_i^{(2)}(b,n) \\ &\quad + \mathbf{V}^{(2)}(b,n) - \mathbf{V}_k^{(1)}(b) \mathbf{X}_k^{(2)}(b,n) \\ &= \mathbf{Y}_k^{(1)}(b) \mathbf{X}_k^{(2)}(b,n) + \tilde{\mathbf{V}}^{(2)}(b,n) \end{aligned} \tag{77}$$

where $\tilde{\mathbf{V}}^{(2)}(b,n) = \mathbf{V}^{(2)}(b,n) + \sum_{i=1,i\neq j}^{K} \mathbf{H}_i^{(2)} \mathbf{S}_i^{(2)}(b,n) + \mathbf{V}^{(2)}(b,n) - \mathbf{V}_k^{(1)}(b) \mathbf{X}_k^{(2)}(b,n)$.

Substituting Eq. (77) into (76), we have

$$\begin{aligned} \mathbf{Y}_k^{(1)} (2\mathbf{I} - (\mathbf{X}_i^{(2)})^H \mathbf{X}_j^{(2)} - \mathbf{X}_i^{(2)}(\mathbf{X}_j^{(2)})^H)(\mathbf{Y}_k^{(1)})^H \\ \leq 2\mathrm{Re}\left\{ (\tilde{\mathbf{V}}^{(2)})^H \mathbf{Y}_k^{(1)}(\mathbf{X}_j^{(2)} - \mathbf{X}_i^{(2)}) \right\}. \end{aligned} \tag{78}$$

Since the identity

$$(\mathbf{X}_i - \mathbf{X}_j)(\mathbf{X}_i - \mathbf{X}_j)^H = 2\mathbf{I} - \mathbf{X}_i^H \mathbf{X}_j - \mathbf{X}_i \mathbf{X}_j^H, \tag{79}$$

Eq. (58) can be simplified to

$$\begin{aligned} \mathbf{Y}_k^{(1)}(\mathbf{X}_i^{(2)} - \mathbf{X}_j^{(2)})(\mathbf{X}_i^{(2)} - \mathbf{X}_j^{(2)})^H (\mathbf{Y}_k^{(1)})^H \\ \leq 2\mathrm{Re}\left\{ (\tilde{\mathbf{V}}^{(2)})^H \mathbf{Y}_k^{(1)}(\mathbf{X}_j^{(2)} - \mathbf{X}_i^{(2)}) \right\}, \end{aligned} \tag{80}$$

which is equal to Eq. (53). Hence, Eq. (52) can be transformed into Eq. (53).

APPENDIX B
COHERENT DETECTOR FOR STBC-SM SCHEME

In coherent STBC-SM scheme, pilots are transmitted in the first sub-block to estimate CSI. After channel estimation (i.e., channel $\mathbf{H}_k \in \mathbb{C}^{N_r \times N_t}$ is known at the BS), the activated antenna combination pair (i_k, j_k) can be estimated by

$$\begin{aligned} (\tilde{i}_1, \tilde{j}_1) &= \underset{\substack{i_k, j_k \in \{[(k-1)N_t, kN_t-1]\} \\ i_k \neq j_k}}{\arg\max} \left| (\mathbf{H}_k\{i_k, j_k\})^H \mathbf{Y}^{(2)}(b,n) \right|, \\ &\quad k = 1, 2, ..., K; n = 1, 2, ..., L/2, \end{aligned} \tag{81}$$

where $\mathbf{H}_k\{i_k, j_k\}$ refers to the i_k-th and j_k-th columns of $\mathbf{H}_k$.

After obtaining the estimated antenna combination pair (i_k, j_k), the transmitted M-PSK symbol pair $(s_{k,1}, s_{k,2})$ can be estimated as

$$(\hat{s}_{k,1}, \hat{s}_{k,2}) = \underset{\substack{\text{coherent} \\ s_{k,1}, s_{k,2}}}{\arg\min} \left\| \mathbf{Y}^{(2)}(b,n) - \mathbf{H}_k \mathbf{A}_k^{(2)}(b,n)\mathbf{\Omega} \right\|_F^2. \tag{82}$$

Next, we will prove that the proposed non-coherent detector performs very close to this coherent detector with perfect CSI when the number of RAs at the BS approaches infinity. Without loss of generality, we only consider the detection

for the 1-st user from $\mathbf{Y}^{(2)}(b,n)$. In coherent detection, it is assume that the channel estimation is error free, the antenna combination pair (i_1, j_1) can be estimated as

$$(\tilde{i}_1, \tilde{j}_1) = \underset{\substack{\text{coherent} \\ i_1, j_1 \in \{[(0, N_t - 1]\} \\ i_1 \neq j_1}}{\arg\max} \left| (\mathbf{H}_1\{i_1, j_1\})^H \mathbf{Y}^{(2)}(b,n) \right|. \tag{83}$$

Since the received signal $\mathbf{Y}^{(2)}(b,n)$ can be reformed as

$$\mathbf{Y}^{(2)}(b,n) = \mathbf{H}_1 \mathbf{S}_1^{(2)}(b,n) + \sum_{k=2}^{K} \mathbf{H}_k \mathbf{S}_k^{(2)}(b,n) + \mathbf{V}^{(2)}(b,n), \tag{84}$$

Eq. (83) can be written as

$$\begin{aligned} (\tilde{i}_1, \tilde{j}_1) &= \underset{\substack{\text{coherent} \\ i_1, j_1 \in \{[(0, N_t - 1]\} \\ i_1 \neq j_1}}{\arg\max} \left| (\mathbf{H}_1\{i_1, j_1\})^H \mathbf{H}_1 \mathbf{S}_1^{(2)}(b,n) \right. \\ &\quad + (\mathbf{H}_1\{i_1, j_1\})^H \sum_{k=2}^{K} \mathbf{H}_k \mathbf{S}_k^{(2)}(b,n) \\ &\quad \left. + (\mathbf{H}_1\{i_1, j_1\})^H \mathbf{V}^{(2)}(b,n) \right|. \end{aligned} \tag{85}$$

Based on the channel asymptotic orthogonality theory in massive MIMO [28], when N_r approaches infinity, Eq. (85) can be simplified as

$$\begin{aligned} (\tilde{i}_1, \tilde{j}_1) &\underset{\text{coherent}}{=}_{N_r \to \infty} \\ &\underset{\substack{i_1, j_1 \in \{[(0, N_t - 1]\} \\ i_1 \neq j_1}}{\arg\max} \left| (\mathbf{H}_1\{i_1, j_1\})^H \mathbf{H}_1 \mathbf{S}_1^{(2)}(b,n) \right|. \end{aligned} \tag{86}$$

If the activated antenna combination pair (i_1, j_1) is estimated correctly by Eq. (83), the transmitted M-PSK symbol pair can be coherently detected as

$$\begin{aligned} (s_{1,1}, s_{1,2}) &= \underset{\substack{\text{coherent} \\ s_{1,1}, s_{1,2}}}{\arg\min} \left\| \mathbf{Y}^{(2)}(b,n) - \mathbf{H}_1 \mathbf{A}_1^{(2)}(b,n)\mathbf{\Omega} \right\|_F^2 \\ &= \underset{s_{1,1}, s_{1,2}}{\arg\min} \left\| \mathbf{Y}^{(2)}(b,n) - \mathbf{H}_1 \left\{ \tilde{i}_1, \tilde{j}_1 \right\} \mathbf{\Omega} \right\|_F^2. \end{aligned} \tag{87}$$

As for our proposed non-coherent scheme, the detection of the activated antenna combination pair for the 1-st user can be written as Eq. (88) at the top of this page. In the sequel, we can obtain Eq. (89) by some basic mathematical calculations.

According to Eqs. (39)-(40) and the fact that the noise matrices are independent of each other and the channel matrix when N_r approaches infinity, Eq. (89) can be simplified as

$$\begin{aligned} (\tilde{i}_1, \tilde{j}_1) &\underset{\text{non-coherent}}{=}_{N_r \to \infty} \\ &\underset{\substack{i_1, j_1 \in \{[(0, N_t - 1]\} \\ i_1 \neq j_1}}{\arg\max} \left| (\mathbf{H}_1\{i_1, j_1\})^H \mathbf{H}_1 \mathbf{X}_1^{(2)}(b,n) \right|. \end{aligned} \tag{90}$$

Based on the law of massive MIMO, it can be conducted that the result of Eq. (90) will converge to that of Eq. (86) when the number of RAs at the BS approaches infinity.

$$\begin{aligned}
(\tilde{i}_1, \tilde{j}_1) &= \underset{\substack{i_1, j_1 \in \{[(0, N_t - 1]\} \\ i_1 \neq j_1}}{\arg\max} \left| (\mathbf{Y}^{(1)}(b)\{i_1, j_1\})^H \mathbf{Y}^{(2)}(b, n) \right| \\
&= \underset{\substack{i_1, j_1 \in \{[(0, N_t - 1]\} \\ i_1 \neq j_1}}{\arg\max} \left| \left(\mathbf{H}\{i_1, j_1\}\mathbf{S}^{(1)}(b) + \mathbf{V}^{(1)}(b) \right)^H \left(\sum_{k=1}^{K} \mathbf{H}_k \mathbf{S}_k^{(2)}(b, n) + \mathbf{V}^{(2)}(b, n) \right) \right| \\
&= \underset{\substack{i_1, j_1 \in \{[(0, N_t - 1]\} \\ i_1 \neq j_1}}{\arg\max} \left| \left(\mathbf{H}_1\{i_1, j_1\}\mathbf{S}_1^{(1)}(b) + \mathbf{V}^{(1)}(b) \right)^H \left(\mathbf{H}_1 \mathbf{S}_1^{(2)}(b, n) + \sum_{k=2}^{K} \mathbf{H}_k \mathbf{S}_k^{(2)}(b, n) + \mathbf{V}^{(2)}(b, n) \right) \right|.
\end{aligned}$$
(88)

$$\begin{aligned}
(\tilde{i}_1, \tilde{j}_1) &= \underset{\substack{i_1, j_1 \in \{[(0, N_t - 1]\} \\ i_1 \neq j_1}}{\arg\max} \left| \left(\mathbf{S}_1^{(1)}(b) \right)^H (\mathbf{H}_1\{i_1, j_1\})^H \mathbf{H}_1 \mathbf{S}_1^{(2)}(b, n) \right. \\
&\quad + \left(\mathbf{S}_1^{(1)}(b) \right)^H (\mathbf{H}_1\{i_1, j_1\})^H \left(\sum_{k=2}^{K} \mathbf{H}_k \mathbf{S}_k^{(2)}(b, n) + \mathbf{V}^{(2)}(b, n) \right) \\
&\quad + \left. \left(\mathbf{V}^{(1)}(b) \right)^H \left(\mathbf{H}_1 \mathbf{S}_1^{(2)}(b, n) + \sum_{k=2}^{K} \mathbf{H}_k \mathbf{S}_k^{(2)}(b, n) + \mathbf{V}^{(2)}(b, n) \right) \right|.
\end{aligned}$$
(89)

Thus, step 1 (i.e., the detection of the activated antenna combination pair in the second sub-block) in the proposed non-coherent detector can hold the same performance as its coherent counterpart.

As for step 2 in the proposed non-coherent detector, the transmitted M-PSK symbol pair $(s_{1,1}, s_{1,2})$ can be estimated by Eq. (91) after estimating the antenna combination pair $(\tilde{i}_1, \tilde{j}_1)$ correctly.

$$\begin{aligned}
(s_{1,1}, s_{1,2}) &= \underset{s_{1,1}, s_{1,2}}{\arg\min} \left\| \mathbf{Y}^{(2)}(b, n) - \mathbf{Y}_1^{(1)}(b) \mathbf{A}_k^{(2)} \boldsymbol{\Omega} \right\|_F^2 \\
&= \underset{s_{1,1}, s_{1,2}}{\arg\min} \left\| \mathbf{Y}^{(2)}(b, n) - \left(\mathbf{Y}_1^{(1)}(b) \right) \{\tilde{i}_1, \tilde{j}_1\} \boldsymbol{\Omega} \right\|_F^2.
\end{aligned}$$
(91)

Since $\left(\mathbf{Y}_1^{(1)}(b) \right) \{\tilde{i}_1, \tilde{j}_1\} = \mathbf{H}_1 \{\tilde{i}_1, \tilde{j}_1\} \mathbf{S}_1^{(1)}(b) + \mathbf{V}_1^{(1)}(b)$ and $\mathbf{S}_1^{(1)}(b)$ is a diagonal matrix with the element of M-PSK symbol, Eq. (91) performs close to its coherent counterpart in Eq. (87) in the high SNR region.

From the derivation of Eqs. (83)-(91), in the high SNR region, the proposed non-coherent detector can provide the same performance as its coherent detector with perfectly channel estimation when the number of RAs tends to infinity.

REFERENCES

[1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.,*, vol. 9, no. 11, pp. 3590-3600, Nov. 2010.

[2] 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.

[3] E. G. Larsson, O. Edfords, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.,*, vol. 52, no. 2, pp. 186-195, Feb. 2014.

[4] A. Chockalingam and B. S. Rajan, Large MIMO systems. Cambridge, U.K.: Cambridge Univ. Press, 2014.

[5] R. 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.

[6] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial modulation for generalized MIMO: challenges, opportunities and implementation," *Proceedings of the IEEE,* vol. 102, no. 1, pp. 26-103, Jan. 2014.

[7] P. Yang, M. Di Renzo, Y. Xiao, S. Li, and L. Hanzo, "Design guidelines for spatial modulation," *IEEE Commun., Surveys Tuts.,* vol. 17, no. 3, pp. 6-26, May. 2014.

[8] T. Lakshmi Narasimhan, P. Raviteja, and A. Chockalingam, "Generalized spatial modulation in large-scale multiuser MIMO systems," *IEEE Trans. Wireless Commun.,* vol. 14, no. 7, pp. 3764-3779, Jul. 2015.

[9] S. Wang, Y. Li, M. Zhang, and J. Wang, "Energy-efficient and low-complexity uplink transceiver for massive spatial modulation MIMO," *IEEE Trans. Veh. Technol.,* vol. 64, no. 10, pp. 4617-4632, Oct. 2015.

[10] L. He, J. Wang, J. Song, and L. Hanzo, "On the multi-user, multi-cell massive spatial modulation uplink: How many antennas for each user?," *IEEE Trans. Wireless Commun.,* vol. 16, no. 3, pp. 1437-1451, Dec. 2016.

[11] Y. Bian, M. Wen, X. Cheng, H. V. Poor, and B. Jiao, "A differential scheme for spatial modulation," in *Proc. 2013 IEEE Globecome. Conf.*, Atlanta, GA, Dec. 2013, pp. 3925-3930.

[12] M. Wen, Z. Ding, X. Cheng, Y. Bian, H. V. Poor, and B. Jiao, "Performance analysis of differential spatial modulation with two transmit antennas," *IEEE Commun. Lett.*, vol. 18, no. 3, pp. 475-478, Mar. 2014.

[13] 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.

[14] J. Li, M. Wen, X. Cheng, Y. Yan, S. Song, and M. H. Lee, "Differential spatial modulation with gray coded antenna activation order," *IEEE Commun. Lett.,* vol. 6, pp. 1100-1103, Jun. 2016.

[15] N. Ishikawa, and S. Sugiura, "Unified differential spatial modulation," *IEEE Wireless Commun. Lett.,* vol. 6, no. 4, pp. 337-340, Feb. 2014.

[16] C. Xu, P. Zhang, R. Rajashekar, N. Ishikawa, S. Sugiura, L. Wang and L. Hanzo, "Finite-cardinality single-RF differential space-time modulation for improving the diversity-throughput tradeoff," *IEEE Trans. Commun.,* vol. 67, no. 1, pp. 318-335, Jan. 2019.

[17] L. Xiao, P. Yang, X. Lei, Y. Xiao, S. Fan, S. Li, and W. Xiang, "A low-complexity detection scheme for differential spatial modulation," *IEEE Commun. Lett.,* vol. 19, no. 9, pp. 1516-1519, Sep. 2015.

[18] N. Ishikawa, R. Rajashekar, C. Xu, M. El-Hajjar, S. Sugiura, L. L. Yang, and L. Hanzo, "Differential-detection aided large-scale generalized spatial modulation is capable of operating in high-mobility millimeter-wave

channels," *IEEE J. Sel. Top. Signal Process.*, vol. 13, no. 6, pp. 1360-1374, Oct. 2019.

[19] L. Xiao, Y. Xiao, P. Yang, J. Liu, S. Li, and W. Xiang, "Space-time block coded differential spatial modulation," *IEEE Trans. Veh. Technol.,* vol. 66, no. 10, pp. 8821-8834, Oct. 2017.

[20] R. Rajashekar, C. Xu, N. Ishikaw, S. Sugiura, K.V.S. Hari, and L. Hanzo, "Full-diversity dispersion matrices from algebraic field extensions for differential spatial modulation," *IEEE Trans. Commun.,* vol. 66, no. 1, pp. 385-394, Jan. 2017.

[21] R. Rajashekar, C. Xu, N. Ishikaw, S. Sugiura, K.V.S. Hari, and L. Hanzo, "Algebraic differential spatial modulation is capable of approaching the performance of its coherent counterpart," *IEEE Trans. Commun.,* vol. 65, no. 10, pp. 4260-4273, Oct. 2017.

[22] N. Ishikawa, and S, Sugiura, "Rectangular differential spatial modulation for open-loop noncoherent massive MIMO downlink," *IEEE Trans. Wireless Commun.,* vol. 16, no. 3, pp. 1908-1920, Mar. 2017.

[23] C. Wu, Y. Xiao, L. Xiao, P. Yang, X. Lei, and W. Xiang, "Space-time block coded rectangular differential spatial modulation: system design and performance analysis," *IEEE Trans. Commun.,* vol. 67, no. 9, pp. 6586-6597, Sep. 2019.

[24] L. Xiao, P. Xiao, H. Ruan, N. Ishikawa, L. Lu, Y. Xiao, and L. Hanzo, "Differentially-encoded rectangular spatial modulation approaches the performance of its coherent counterpart," *IEEE Trans. Commun.,* 2020. DOI:10.1109/TCOMM.2020.3021117.

[25] Y. Jia, S. Li, J. Zhang, N. Wang, and X. Mu, "PSK modulated differential spatial modulation for massive MIMO uplink system," in *Proc. of IET ICWMMN 2017* Beijing, China, Nov. 17-20, 2017, pp. 83-87.

[26] R. Khanzadeh, M. F. Naeiny, "Multiuser differential spatial modulation scheme for massive MIMO uplink communications," *Physical Commun.,* vol. 41, pp. 1-10, May. 2020.

[27] 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.

[28] A. M. Tulino, and S. Verdu, "Random matrix theory and wireless communications," *Found. Trends Commun. Inf. Theory,* vol. 1, no. 1, pp. 1-182, 2004.

[29] M. S. Alaouni, and M. K. Simon, "Digital communication over fading channels," 2nd ed. Hoboken, NJ, USA: Wiley, 2005.

[30] B. D. O. Anderson, and J. B. Moore, "Optimal Filtering", New York: Dover, 2005.

[31] G. L. Turin, "The characteristic function of hermitian quadratic forms in complex normal random variables," *Biometrika,* vol. 47, no. 1/2, pp. 199-201, June. 1960.

[32] W. H. Sheen, and G. L. Stuber, "MLSE equalization and decoding for multipath-fading channels," *IEEE Trans. Commun.,* vol. 39, no. 10, pp. 1455-1464, Oct. 1991.

[33] V. Tarokh, and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE J. Sel. Areas Commun.,* vol. 18, no. 7, pp. 1169-1174, Jul. 2000.

[34] R. Rajashekar and K. V. S. Hari, "Modulation diversity for spatial modulation using complex interleaved orthogonal design," in *Proc. TENCON.,* pp. 1-6, 2012.

[35] L. Xiao, D. Chen, I. Hemadeh, P. Xiao, and T. Jiang, "Generalized space time block coded spatial modulation for open-loop massive MIMO downlink communication systems," *IEEE Trans. Commun.,* vol. 68, no. 11, pp. 6858-6871, Nov. 2020.

[36] T. Marzetta, "How much training is required for multiuser MIMO?" in *Proc. 40th Asilomar Conf. Signals Syst. Comput. (ACSSC)*, Oct/Nov. 2006, pp. 359-363.

[37] A. Chockalingam, and B. S. Rajan, *Large scale MIMO systems*, Cambridge, U.K.: Cambridge Univ. Press, Feb. 2014.