

Received February 24, 2017, accepted March 19, 2017, date of publication March 27, 2017, date of current version April 24, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2688019

NOMA-Based Spatial Modulation

XUDONG ZHU¹, ZHAOCHENG WANG¹, AND JIANFEI CAO²

¹Tsinghua National Laboratory for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

²SONY China Research Laboratory, Beijing 100084, China

Corresponding author: Z. Wang (zawang@tsinghua.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61571267, in part by the Shenzhen Subject Arrangements under Grant JCYJ20160331184124954, in part by the Shenzhen Fundamental Research Project under Grant JCYJ20150401112337177, and in part by the Sony Corporation and Sony China Research Laboratory, Sony (China) Ltd.

ABSTRACT Spatial modulation (SM) has emerged as a low-complexity and energy-efficient multiple-input multiple-output transmission technique, where the information bits are not only transmitted by amplitude phase modulation but also conveyed by the index of activated transmit antenna (TA). By deploying SM in downlink multi-user (DL-MU) scenarios, conventional orthogonal multiple access-based SM (OMA-SM) allocates exclusive time-frequency resources to users, but suffers from low spectral efficiency. TA grouping-based SM (TAG-SM) divides TAs into sub-groups to serve different users independently, but suffers from severe inter-user interference. By introducing non-OMA (NOMA) into SM for DL-MU transmission, NOMA-based SM (NOMA-SM) is proposed to mitigate inter-user interference, while maintaining high spectral efficiency. Specifically, by applying successive interference cancellation at user side, the inter-user interference could be effectively eliminated with the sacrifice of increased computational complexity. Afterward, based on a symbol error rate analysis, a low-complexity power allocation scheme is provided to achieve high spectral efficiency through power domain multiplexing. When considering near-far effect from user distribution, user pairing issue is also discussed. Numerical simulation compares NOMA-SM with OMA-SM and TAG-SM, and verifies the effectiveness of the proposed low-complexity power allocation and user pairing methodologies.

INDEX TERMS Spatial modulation (SM), non-orthogonal multiple access (NOMA), successive interference cancellation (SIC), multiple-input multiple-output system (MIMO), downlink multiuser (DL-MU).

I. INTRODUCTION

To meet the escalating demand of wireless data traffic by smart phones, tablets and various applications such as live web cast and cloud computing, the cellular network has been densified to improve the spectral efficiency [1], [2]. At the same time, considering the requirement to curtail the carbon emission, energy reduction of base stations (BSs) in future 5G cellular network is of great importance [3], [4]. A typical BS for cellular network consists of baseband processors, radio-frequency (RF) chains and transmit antennas (TAs), where energy consumption mainly comes from the power amplifiers (PAs), which are equipped in each RF chain [5]. The energy consumption challenges the application of conventional MIMO schemes, which requires the same number of RF chains as TAs [6], [7].

Spatial modulation (SM) has emerged as a low-complexity and energy-efficient MIMO transmission technique for future 5G cellular network [8]–[12]. Specifically, SM is a

point-to-point single-stream transmission scheme, where the bit stream is divided into groups and each group consists of two parts: 1) the first part of bit information is mapped to the index of activated TA to transmit a signal while other TAs remain silent; 2) the second part of bit information is mapped to a symbol from the conventional amplitude and phase modulation (APM) constellation diagram, which will be radiated through the activated TA. Compared with the conventional MIMO schemes which require the same number of RF chains as TAs, SM enjoys high energy efficiency due to the fact that only one or a small number of RF chains are required, regardless of the number of TAs [8]–[12]. By introducing the additional transmission channel by TA index, the spatial multiplexing gain of SM schemes is comparable to the conventional multi-stream MIMO schemes [13], [14].

The classical SM schemes, such as space shift keying (SSK) [8], generalized SM (GSM) [9], [10], trellis coded SM (TC-SM) [11] and their variants are all devoted to

a point-to-point MIMO transmission scenario. Since a BS typically serves multi-user (MU) within the cell, classical SM schemes should evolve to support MU transmission for cellular networks [1]–[4]. For uplink MU (UL-MU) transmission, each user adopts SM independently and the BS with multiple receive antennas (RAs) could separate the signals from different users at the cost of increased detection complexity [15]–[17]. Several strategies have been proposed to achieve a trade-off between detection performance and computation complexity, e.g., message passing based low-complexity detection algorithm [15], massive MIMO with low-resolution analog-to-digital convertors (ADCs) [16] and compressive sensing (CS) aided detection scheme [17]. For downlink MU (DL-MU) transmission, conventional orthogonal multiple access (OMA) techniques, e.g., time division multiple access (TDMA) and frequency division multiple access (FDMA), could be directly combined with SM schemes (named OMA-SM), but suffer from low spectral efficiency due to exclusive time-frequency resource allocation. Besides division over time and frequency domains, TA domain division could also be utilized for multiple access [18]–[20]. Specifically, TAs at the BS are divided into several groups and each group of TAs are utilized to implement SM transmission for different users independently. Different from user orthogonal property over time and frequency domains, such TA grouping based SM (TAG-SM) schemes suffer from severe inter-user interference. Precoding schemes based on power allocation are proposed to suppress such inter-user interference by assuming channel state information (CSI) at the BS [18], [19]. By introducing adaptive TA selection instead of fixed TA grouping procedure, spectral efficiency could be further improved [20], [21]. Moreover, a layered SM scheme for DL-MU transmission is proposed in [22], where user bit streams are combined and then transmitted in a broadcast way with low spectral efficiency. Therefore, inter-user interference and spectral efficiency become two bottlenecks in DL-MU transmission, while existing SM based DL-MU transmission schemes could only achieve one issue, whereas suffer from the other [18]–[22].

To achieve higher spectral efficiency for 5G cellular networks, a non-orthogonal multiple access (NOMA) over power domain has been proposed as a promising solution [23], [24]. Unlike traditional OMA techniques, where exclusive time-frequency resource are allocated, NOMA integrates different users' data in the power domain. At user side, successive interference cancellation (SIC) should be deployed to remove the inter-user interference before data detection [23]. The NOMA performance for DL-MU transmission in a cellular network has been analytically investigated through the outage probability and ergodic sum rate, which has verified the performance gain by NOMA over its traditional OMA counterparts [25]. The power allocation problem among users for NOMA is discussed in [26], where users with poor channel conditions are allocated with more power to detect their signal directly while users with favorable channel conditions are allocated

with less power to facilitate the SIC procedure. The impact of user pairing is studied in [27], which indicates that users with more distinctive channel conditions should be paired, which could achieve higher spectral efficiency.

In this paper, by introducing NOMA [23], [24] into DL-MU transmission combined with SM methodology, NOMA-SM is proposed to achieve high spectral efficiency for 5G cellular networks, meanwhile inter-user interference is handled by SIC procedure at user side. Different from neither OMA-SM nor TAG-SM [18]–[20], where exclusive time-frequency or TA resource are allocated to different users, NOMA-SM utilizes the same time-frequency and TAs to serve MU with carefully designed power allocation. The main outcome and contribution of our paper are summarized as follows.

- *Overview of OMA-SM and TAG-SM:* A brief overview of OMA-SM and TAG-SM [18]–[20] is provided as counterparts of the proposed NOMA-SM scheme. OMA-SM allocates exclusive time-frequency resource to different users, but suffers from low spectral efficiency. TAG-SM divides TAs into sub-groups to serve users independently, which is able to improve the spectral efficiency compared to OMA-SM. However, inter-user interference is introduced due to TA domain division, which leads to a detection error floor at user side.
- *NOMA-SM Design:* By combining NOMA with SM for DL MU transmission, the proposed NOMA-SM design is able to improve the spectral efficiency compared to TAG-SM. Signals intended to different users are integrated at power domain and then radiated through the same time-frequency and TA resource. By deploying SIC based detection scheme, inter-user interference could be mitigated at the cost of increased computational complexity. Compared with OMA-SM and TAG-SM, NOMA-SM achieves a better trade-off between spectral efficiency and interference mitigation.
- *Performance Analysis and Practical Considerations:* Firstly, theoretical symbol error rate (SER) analysis is presented, which indicates the error propagation issue among users due to the SIC procedure. Afterwards, the power allocation issue is formulated as optimization problems to minimize the maximum SER or minimize the average SER. Based on the SER analysis for NOMA-SM, a low-complexity power allocation scheme is proposed to achieve high spectral efficiency through power domain multiplexing. Moreover, by considering near-far effect from user distribution, user pairing is also discussed.

In simulation, theoretical SER analysis is firstly verified. After that, we demonstrate that the proposed low-complexity power allocation scheme is able to provide acceptable performance compared to simulation results. BER performance comparison between OMA-SM, TAG-SM and the proposed NOMA-SM is presented in detail.

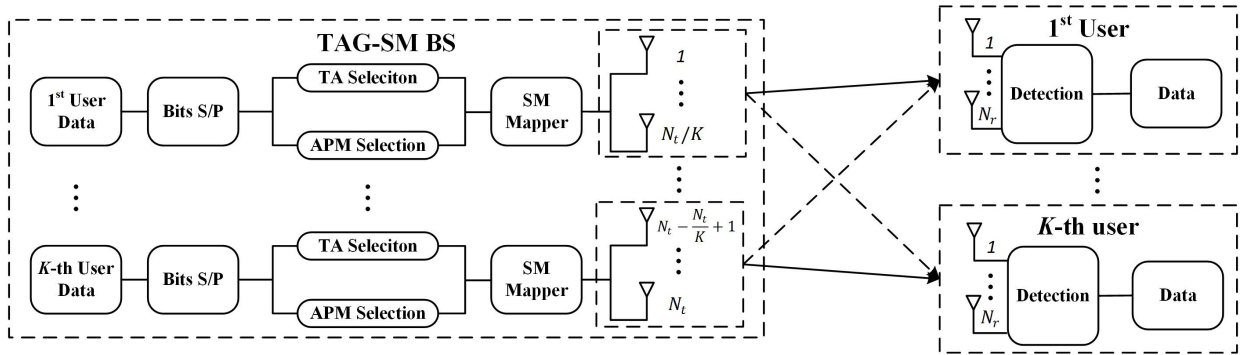


FIGURE 1. Transceiver architecture of TAG-SM.

The rest of the paper is organized as follows. In Section II, a brief overview of OMA-SM and TAG-SM for DL-MU transmission is presented as counterparts of NOMA-SM. Section III details the principle of NOMA-SM, including transmitter, receiver and its performance comparison with OMA-SM and TAG-SM. Section IV provides the performance analysis of NOMA-SM over SER approximation, power allocation and user pairing. Our simulation results indicating the benefits of NOMA-SM are presented in Section V, while our conclusions are drawn in Section VI.

Throughout this paper, boldface lower and upper-case symbols represent vectors and matrices, respectively. We use $(\cdot)^T$ for transpose, $(\cdot)^*$ for complex conjugate, $(\cdot)^H$ for conjugate transpose, $\binom{\cdot}{\cdot}$ for the binomial coefficient and $\|\cdot\|_p$ for the p -th order norm of a vector/matrix. We use $\mathcal{CN}(\mu, \sigma^2)$ for the complex Gaussian distribution of a random variable, having independent Gaussian distribution real and imaginary parts denoted by $\mathcal{N}(\mu, \frac{\sigma^2}{2})$, with mean μ and variance $\frac{\sigma^2}{2}$. We use $P(\cdot)$ for the probability of an event and $\mathbb{E}_x\{\cdot\}$ for the statistical expectation with respect to x . We use $\log(\cdot)$ for the natural logarithm and $\log_a(\cdot)$ for the logarithm with base a .

II. OVERVIEW OF OMA-SM AND TAG-SM

A. OMA-SM FOR DL-MU TRANSMISSION

In this paper, DL-MU transmission is considered, where the BS is equipped with N_t TAs to serve K users [1], [2]. We assume that all users have N_r RAs to receive the DL signal. Rayleigh fading channel between the BS and the k -th user ($1 \leq k \leq K$) [18]–[20] is considered as

$$\mathbf{H}^{(k)} = [\mathbf{h}_1^{(k)} \mathbf{h}_2^{(k)} \dots \mathbf{h}_{N_t}^{(k)}] \in \mathbb{C}^{N_r \times N_t}, \quad (1)$$

where $\mathbf{h}_j^{(k)} = [h_{1,j}^{(k)} h_{2,j}^{(k)} \dots h_{N_r,j}^{(k)}]^T$ denotes the channel gain between the j -th TA at BS and the RAs at the k -th user with independent and identical distribution (i.i.d.) as $\mathbf{h}_{i,j}^{(k)} \sim \mathcal{CN}(0, 1)$.

By adopting OMA techniques, the SM for DL-MU transmission could be simplified as a point-to-point transmission. As an example, by combining TDMA with SM for DL-MU transmission, the data rate (bits/symbol) of the

k -th user ($1 \leq k \leq K$) could be represented as

$$R_{\text{OMA-SM},k} = \frac{1}{K} \left(\log_2(N_t) + \log_2(M) \right), \quad (2)$$

where $\log_2(N_t)$ denotes the information bits conveyed by the activated TA index (here we assume only one TA is activated at any time slot), $M = |\mathcal{X}|$ denotes the number of symbols in an APM diagram \mathcal{X} , such as phase shift keying (PSK) or quadrature amplitude modulation (QAM) and coefficient $\frac{1}{K}$ denotes that exclusive time-frequency resource is equally allocated to K users without inter-user interference.

Although OMA-SM schemes enjoy free of inter-user interference, they suffer from spectral efficiency loss due to their inherent exclusive time-frequency resource allocation. When considering TDMA combined with SM, only one RF chain is required at the BS and hence the hardware cost of MU scenario is almost the same as the single-user (SU) scenario.

B. TAG-SM FOR DL-MU TRANSMISSION

Different from OMA-SM where exclusive time-frequency resource is allocated for various users, TAG-SM regards TAs at the BS as resource in spatial dimension to be allocated to different users [18]–[20]. As shown in Fig. 1, N_t TAs at the BS could be directly divided into K subgroups [18] (assuming $\frac{N_t}{K}$ is an integer for simplicity), e.g., the k -th TA subgroup serving the k -th user consists of TAs with indices in the range $[(k-1)\frac{N_t}{K} + 1, k\frac{N_t}{K}]$. Hence, for the k -th user, only channel vectors $\mathbf{h}_j^{(k)}$, $(k-1)\frac{N_t}{K} + 1 \leq j \leq k\frac{N_t}{K}$ are valid to transmit the APM symbol.

For DL bit stream of the k -th user, a random sequence of independent bits $\mathbf{b}^{(k)} = [b_1^{(k)} b_2^{(k)} \dots]$ enters the serial/parallel (S/P) converter and outputs two bit streams for TA selection and APM selection as shown in Fig. 1. After SM mapper, the signal vector for the k -th TA subgroup (the k -th user's data) could be represented as

$$\mathbf{x}_{j_k, m_k}^{(k)} \triangleq \sqrt{\frac{\rho}{K}} \begin{bmatrix} 0 & \dots & 0 & s_{m_k} & 0 & \dots & 0 \end{bmatrix}^T \in \mathbb{C}^{\frac{N_t}{K} \times 1}, \quad (3)$$

↑
 j_k -th position

where $1 \leq j_k \leq \frac{N_t}{K}$ denotes the selected TA index in the k -th TA subgroup, s_{m_k} ($1 \leq m_k \leq M$) denotes the

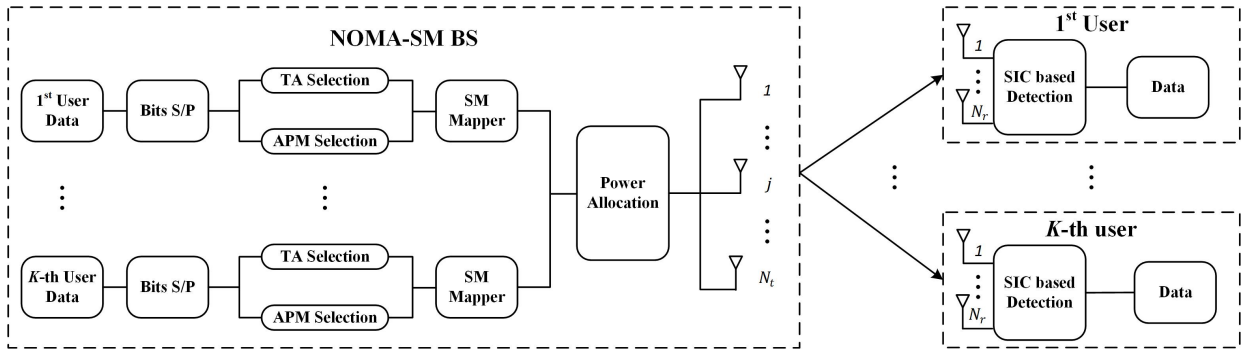


FIGURE 2. Transceiver architecture of the proposed NOMA-SM.

m_k -th symbol from the M -ary APM constellation diagram with $\mathbb{E}\{s_i^* s_i\} = 1$ and ρ denotes the total transmit power for all K users. Hence, the data rate (bits/symbol) of the k -th user in TAG-SM could be represented as

$$R_{\text{TAG-SM},k} = \log_2\left(\frac{N_t}{K}\right) + \log_2(M). \quad (4)$$

By adopting the same APM order, the data rate of TAG-SM is higher than OMA-SM, but at the cost of severe inter-user interference.

Through the wireless channel $\mathbf{H}^{(k)}$, the received signal at the k -th user is given by

$$\begin{aligned} \mathbf{y}^{(k)} &= \mathbf{H}^{(k)} \left[(\mathbf{x}_{j_1, m_1}^{(1)})^T \cdots (\mathbf{x}_{j_K, m_K}^{(K)})^T \right]^T + \mathbf{n}^{(k)} \\ &= \sqrt{\frac{\rho}{K}} (\mathbf{h}_{(k-1)\frac{N_t}{K} + j_k}^{(k)} s_{m_k} + \sum_{l \neq k} \mathbf{h}_{(l-1)\frac{N_t}{K} + j_l}^{(k)} s_{m_l}) + \mathbf{n}^{(k)}, \end{aligned} \quad (5)$$

where $\sum_{l \neq k} \mathbf{h}_{(l-1)\frac{N_t}{K} + j_l}^{(k)} s_{m_l}$ denotes the inter-user interference caused by other users and $\mathbf{n}^{(k)} = [n_1^{(k)} \ n_2^{(k)} \ \cdots \ n_{N_r}^{(k)}]^T$ denotes the N_r -dimensional additive white Gaussian noise (AWGN) with i.i.d. entries as $n_i^{(k)} \sim \mathcal{CN}(0, \sigma_n^2)$.

From Eq. (5), the effect of AWGN $\mathbf{n}^{(k)}$ could be reduced by increasing the total transmit power ρ , while the inter-user interference $\sum_{l \neq k} \mathbf{h}_{(l-1)\frac{N_t}{K} + j_l}^{(k)} s_{m_l}$ becomes the performance bottleneck. Hence, without inter-user interference cancellation, there exists an error floor in data detection at user side. Adaptive TA selection and precoding have been proposed to suppress this inter-user interference [18]–[20], but CSI is required at the BS which introduces additional training overhead. Moreover, compared to OMA-SM which only requires one RF chain, K RF chains are needed at the transmitter when TAG-SM is deployed.

III. PROPOSED NOMA-SM

To improve spectral efficiency and mitigate inter-user interference at the same time, we combine NOMA with SM for DL-MU transmission. In this section, transmitter and receiver of NOMA-SM as well as performance comparison with OMA-SM and TAG-SM are presented in detail.

A. NOMA-SM TRANSMITTER

The transceiver architecture of NOMA-SM is illustrated in Fig. 2. Different from TAG-SM which divides TAs into subgroups to serve each user independently, all users in NOMA-SM share N_t TAs at the BS for DL transmission with carefully designed power allocation. Specifically, for the k -th user, the signal vector transmitted over N_t TAs could be represented as

$$\tilde{\mathbf{x}}_{j_k, m_k}^{(k)} \triangleq \sqrt{\rho_k} [0 \cdots 0 \ s_{m_k} \ 0 \cdots 0]^T \in \mathbb{C}^{N_t \times 1}, \quad (6)$$

\uparrow
 j_k -th position

where $1 \leq j_k \leq N_t$ denotes the selected TA index among N_t TAs, s_{m_k} ($1 \leq m_k \leq M$) denotes the m_k -th symbol from the M -ary APM constellations and ρ_k denotes the transmit power allocated to the k -th user with the total power constraint $\sum_{k=1}^K \rho_k = \rho$. PSK modulation is applied and we have $s_i^* s_i = 1$, $1 \leq i \leq M$.

The data rate of the k -th user in the proposed NOMA-SM could be represented as

$$R_{\text{NOMA-SM},k} = \log_2(N_t) + \log_2(M). \quad (7)$$

Compared with OMA-SM and TAG-SM as shown in Eq. (2) and (4), the data rate of the proposed NOMA-SM could be improved significantly.

B. NOMA-SM RECEIVER

In OMA-SM and TAG-SM, the conventional SM detector in SU scenario could be directly deployed in DL-MU transmission. However, as shown in Fig. 2, a complicated SIC procedure has to be deployed at user side in NOMA-SM to mitigate inter-user interference before data detection, which leads to increased computational complexity.

The received DL signal at the k -th user in NOMA-SM could be represented as

$$\begin{aligned} \tilde{\mathbf{y}}^{(k)} &= \mathbf{H}^{(k)} \sum_{l=1}^K \tilde{\mathbf{x}}_{j_l, m_l}^{(l)} + \mathbf{n}^{(k)} \\ &= \sqrt{\rho_k} \mathbf{h}_{j_k}^{(k)} s_{m_k} + \sum_{l \neq k} \sqrt{\rho_l} \mathbf{h}_{j_l}^{(k)} s_{m_l} + \mathbf{n}^{(k)}. \end{aligned} \quad (8)$$

Since the signal received at the k -th user is disturbed by other users' data $\sum_{l \neq k} \sqrt{\rho_l} \mathbf{h}_{j_l}^{(k)} s_{m_l}$, SIC procedure has to be

combined with NOMA-SM at user side to mitigate inter-user interference [23], [24]. The basic idea of SIC is to remove the data intended for other users with larger transmit power. Without loss of generality, power allocation is designed in a descending order with total power constraint as

$$\rho_1 > \rho_2 > \dots > \rho_K, \sum_{k=1}^K \rho_k = \rho. \quad (9)$$

Since the 1st user has the highest transmit power, the SIC procedure is unnecessary and its data could be directly detected by the conventional SM based detection methods, e.g., maximum likelihood (ML) [12]

$$[\hat{j}_1, \hat{m}_1] = \arg \max_{j,m} \|\tilde{\mathbf{y}}^{(1)} - \sqrt{\rho_1} \mathbf{h}_j^{(1)} s_m\|_2^2, \quad (10)$$

where $[\hat{j}_1, \hat{m}_1]$ denotes the detected data at the 1st user side. However, for the k -th user ($k > 1$), data detection becomes much more complicated compared to the 1st user since it has to remove the interference from the previous $k - 1$ users in a sequential way [23], [24]. After that, the updated received signal at the k -th user could be represented as

$$\tilde{\mathbf{y}}_{k-1}^{(k)} = \tilde{\mathbf{y}}^{(k)} - \sum_{l=1}^{k-1} \sqrt{\rho_l} \mathbf{h}_{j_l,k}^{(k)} s_{\hat{m}_l,k}, \quad (11)$$

where $[\hat{j}_{l,k}, \hat{m}_{l,k}]$ with $1 \leq l \leq k - 1$ denotes the detected data of the l -th user, which should be removed at the k -th user. After the SIC procedure as shown in Eq. (11), the k -th user is able to detect its own data as

$$[\hat{j}_k, \hat{m}_k] = \arg \max_{j,m} \|\tilde{\mathbf{y}}_{k-1}^{(k)} - \sqrt{\rho_k} \mathbf{h}_j^{(k)} s_m\|_2^2. \quad (12)$$

Moreover, similar detection procedure could be easily obtained by deploying other detection algorithms, e.g., maximum ratio combining (MRC) [12], vector based detection (VBD) [28] and so on.

TABLE 1. Comparison from system aspects.

	OMA-SM	TAG-SM	NOMA-SM
Spectral efficiency	Low	Middle	High
Inter-user interference	No	Severe	Acceptable
Encoder/decoder complexity	Low	Middle	High

C. COMPARISON OF OMA-SM, TAG-SM AND NOMA-SM

As shown in Table 1, we compare the proposed NOMA-SM with OMA-SM and TAG-SM in detail as follows.

- *Spectral efficiency*: The data rate of OMA-SM is much lower than TAG-SM and NOMA-SM due to exclusive time-frequency resource allocation. Compared with TAG-SM, NOMA-SM enjoys an additional data rate as $R_{\text{NOMA-SM},k} - R_{\text{TAG-SM},k} = \log_2(K)$, since all users could utilize N_t TAs for DL transmission.
- *Inter-user Interference*: By allocating exclusive time-frequency resource to different users, OMA-SM almost has no inter-user interference. While in TAG-SM, inter-user interference mitigation schemes have to

be deployed at the BS side [18]–[20], where CSI requirement leads to additional training overhead. For NOMA-SM, the inter-user interference could be mitigated at user side by deploying SIC procedure, the computational complexity is increased by K times at most compared with OMA-SM and TAG-SM.

- *Encoder/decoder complexity*: Considering TDMA with OMA-SM, only one RF chain is required to support MU-DL transmission. However, both TAG-SM and NOMA-SM require K RF chains at the transmitter to support K data streams simultaneously. Moreover, NOMA-SM also requires power allocation module at the BS and SIC module at user side.

As a whole, NOMA-SM is able to achieve a better trade-off between data rate and interference mitigation in DL-MU transmission with the sacrifice of moderately increased computational complexity and hardware complexity.

IV. PERFORMANCE ANALYSIS AND PRACTICAL CONSIDERATIONS

In this section, performance analysis and practical considerations for NOMA-SM are presented, including SER analysis, low-complexity power allocation and user pairing for practical implementation.

A. SER ANALYSIS

Based on Eq. (8), the equivalent interference to the 1st user could be represented as

$$\mathbf{I}^{(1)} = \sum_{l=2}^K \sqrt{\rho_l} \mathbf{h}_{j_l}^{(1)} s_{m_l} + \mathbf{n}^{(1)}, \quad (13)$$

where $\mathbf{h}_{j_l}^{(1)} = [h_{1,j_l}^{(1)} \ h_{2,j_l}^{(1)} \ \dots \ h_{N_r,j_l}^{(1)}]^T$, $h_{i,j_l}^{(1)} \sim \mathcal{CN}(0, 1)$, $|s_{m_l}| = 1$, $1 \leq j_l \leq N_t$ and $\mathbf{n}^{(1)} = [n_1^{(1)} \ n_2^{(1)} \ \dots \ n_{N_r}^{(1)}]^T$ with $n_i^{(1)} \sim \mathcal{CN}(0, \sigma_n^2)$. Hence, the element in the equivalent interference vector $\mathbf{I}^{(1)}$ to the 1st user has the following i.i.d. distribution

$$I_i^{(1)} \sim \mathcal{CN}(0, \sum_{l=2}^K \rho_l + \sigma_n^2), \quad (14)$$

where $I_i^{(1)}$, $1 \leq i \leq N_r$ denotes the i -th element in $\mathbf{I}^{(1)}$. The symbol error rate (SER) and bit error rate (BER) for classical SM in SU scenario have been widely studied [12]–[14]. Here a function f_{SER} is defined to represent the SER of a point-to-point M -ary PSK based SM transmission with a $N_t \times N_r$ Rayleigh fading channel as $f_{\text{SER}}(\frac{\rho}{\sigma_n^2})$, where $\frac{\rho}{\sigma_n^2}$ denotes the average signal-to-noise-ratio (SNR) at the RAs. Various approximate expressions and bounds could be found about this function $f_{\text{SER}}(\frac{\rho}{\sigma_n^2})$ [12]–[14]. Hence, the SER of the 1st user in NOMA-SM could be approximated as

$$\text{SER}_{\text{NOMA-SM},1} \stackrel{(a)}{\approx} f_{\text{SER}}\left(\frac{\rho_1}{\sum_{l=2}^K \rho_l + \sigma_n^2}\right), \quad (15)$$

where $\stackrel{(a)}{\approx}$ denotes that $\mathbf{I}^{(1)}$ is not independent of $\sqrt{\rho_1} \mathbf{h}_{j_1}^{(1)} s_{m_1}$ (TA selected by the 1st user could also be selected by the

$$\text{SER}_{\text{NOMA-SM},k} = P_{k,1}(\hat{j}_k, \hat{m}_k \neq [j_k, m_k], \exists 1 \leq l \leq k-1, \text{s.t.}, [\hat{j}_{l,k}, \hat{m}_{l,k}] \neq [j_l, m_l]) \\ + P_{k,2}(\hat{j}_k, \hat{m}_k \neq [j_k, m_k], [\hat{j}_{l,k}, \hat{m}_{l,k}] = [j_l, m_l], \forall 1 \leq l \leq k-1). \quad (17)$$

$$P_{k,1} \stackrel{(b)}{\approx} 1 - P([\hat{j}_{l,k}, \hat{m}_{l,k}] = [j_l, m_l], \forall 1 \leq l \leq k-1) \stackrel{(a)}{\approx} 1 - \prod_{l=1}^{k-1} \left(1 - f_{\text{SER}}\left(\frac{\rho_l}{\sum_{t=l+1}^K \rho_t + \sigma_n^2}\right)\right) \\ \stackrel{(c)}{\approx} \sum_{l=1}^{k-1} f_{\text{SER}}\left(\frac{\rho_l}{\sum_{t=l+1}^K \rho_t + \sigma_n^2}\right). \quad (18)$$

$$P_{k,2} \stackrel{(a)}{\approx} f_{\text{SER}}\left(\frac{\rho_k}{\sum_{t=k+1}^K \rho_t + \sigma_n^2}\right) * P([\hat{j}_{l,k}, \hat{m}_{l,k}] = [j_l, m_l], \forall 1 \leq l \leq k-1) \\ \stackrel{(b)}{\approx} f_{\text{SER}}\left(\frac{\rho_k}{\sum_{t=k+1}^K \rho_t + \sigma_n^2}\right) * \left(1 - \sum_{l=1}^{k-1} f_{\text{SER}}\left(\frac{\rho_l}{\sum_{t=l+1}^K \rho_t + \sigma_n^2}\right)\right) \\ \stackrel{(c)}{\approx} f_{\text{SER}}\left(\frac{\rho_k}{\sum_{t=k+1}^K \rho_t + \sigma_n^2}\right). \quad (19)$$

latter users). After that, by considering SIC based ML detector at the k -th user, we have the following corollary about the SER approximation of the k -th user.

Corollary 1: The SER of the k -th user with SIC based ML detector in NOMA-SM transmission could be approximated at high SNR region as

$$\text{SER}_{\text{NOMA-SM},k} \approx \sum_{l=1}^k f_{\text{SER}}\left(\frac{\rho_l}{\sum_{t=l+1}^K \rho_t + \sigma_n^2}\right). \quad (16)$$

Proof: Firstly, the SER of the k -th user is divided into two parts as shown in Eq. (17), as shown at the top of this page. The first part $P_{k,1}$ denotes the SER when SIC procedure is correctly carried out, while the second part $P_{k,2}$ denotes the SER when any errors occur in SIC procedure. $P_{k,1}$ is directly approximated by the SIC failure probability in Eq. (18), as shown at the top of this page, where $\stackrel{(b)}{\approx}$ indicates that the tiny probability that the k -th user correctly detects its data when errors occur in the SIC procedure is ignored. After that, $\stackrel{(c)}{\approx}$ indicates that the quadratic terms and high-order terms of f_{SER} in $\prod_{l=1}^{k-1} \left(1 - f_{\text{SER}}\left(\frac{\rho_l}{\sum_{t=l+1}^K \rho_t + \sigma_n^2}\right)\right)$ could be ignored when f_{SER} approaches to zero at high SNR region. Similar to $P_{k,1}$, $P_{k,2}$ could also be approximated by f_{SER} function as shown in Eq. (19), as shown at the top of this page. Finally, the SER of the k -th user with SIC based ML detector in NOMA-SM could be approximated by Eq. (16) at high SNR region.

Based on *Corollary 1*, the detection error of the previous users will be propagated to the latter users. Hence, the latter users in NOMA-SM suffer from error propagation, which is unavoidable due to the SIC procedure. For practical implementation, the number of users in NOMA-SM should be kept small to avoid performance degradation caused by error propagation, e.g., typically $K = 2$ users.

B. LOW-COMPLEXITY POWER ALLOCATION

For practical implementation, power allocation should be carefully designed to achieve performance balance among all K users. By considering different optimization targets, we could formulate different problems, e.g., minimize the maximal SER of K users, minimize the average SER of all K users and so on.

Firstly, we formulate the optimization problem P_1 to minimize the maximal SER of K users as

$$P_1 : \{\rho_k^{\text{opt},1}\} = \arg \min_{\{\rho_k\}} \max \{\text{SER}_{\text{NOMA-SM},k} : 1 \leq k \leq K\} \\ \text{s.t. } \rho_1 > \rho_2 > \dots > \rho_K, \sum_{k=1}^K \rho_k = \rho. \quad (20)$$

According to *Corollary 1*, the SER of the K -th user tends to be much higher than the previous $K - 1$ users. Hence, the optimization problem P_1 to minimize the maximal SER of K users could be approximated by minimizing the SER of the K -th user as

$$P'_1 : \{\rho_k^{\text{opt},1'}\} = \arg \min_{\{\rho_k\}} \text{SER}_{\text{NOMA-SM},K} \\ \approx \arg \min_{\{\rho_k\}} \sum_{l=1}^K f_{\text{SER}}\left(\frac{\rho_l}{\sum_{t=l+1}^K \rho_t + \sigma_n^2}\right), \quad (21)$$

where the power constraint is omitted here for simplicity. By assuming the prior knowledge of function f_{SER} [12]–[14] and a small number of users, e.g., $K = 2$, the problem P'_1 could be easily solved by the derivation of ρ_1 ($\rho_2 = \rho - \rho_1$) or grid search [29].

Secondly, to minimize the average SER of all K users, the power allocation problem could be formulated as

$$P_2 : \{\rho_k^{\text{opt},2}\} = \arg \min_{\{\rho_k\}} \overline{\text{SER}} \\ \text{s.t. } \rho_1 > \rho_2 > \dots > \rho_K, \sum_{k=1}^K \rho_k = \rho, \quad (22)$$

$$\begin{aligned} \overline{\text{SER}} &= \frac{1}{K} \sum_{k=1}^K \text{SER}_{\text{NOMA-SM},k} \approx \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^k f_{\text{SER}}\left(\frac{\rho_l}{\sum_{t=l+1}^K \rho_t + \sigma_n^2}\right) \\ &\approx f_{\text{SER}}\left(\frac{\rho_1}{\sum_{t=2}^K \rho_t + \sigma_n^2}\right) + \frac{K-1}{K} f_{\text{SER}}\left(\frac{\rho_2}{\sum_{t=3}^K \rho_t + \sigma_n^2}\right) + \dots + \frac{1}{K} f_{\text{SER}}\left(\frac{\rho_K}{\sigma_n^2}\right). \end{aligned} \quad (23)$$

where detailed expression of $\overline{\text{SER}}$ is presented in Eq. (23), as shown at the top of this page. Similar to P'_1 , P_2 could also be solved with a small number of users [23]–[25], e.g., $K = 2, 3$.

For practical implementation, a low-complexity power allocation scheme is proposed to provide numerical solutions to P'_1 and P_2 .

- **Power division:** The total power ρ is uniformly divided into N parts, and the power allocation ρ_k is constrained to be an integral multiple of $\frac{\rho}{N}$, i.e., $\rho_k = T_k \frac{\rho}{N}$ (T_k is assumed to be an integer).
- **Power allocation patterns:** The original power allocation constraints could be reformulated as $T_1 > T_2 > \dots > T_K$, $\sum_{k=1}^K T_k = N$. Since T_k should be an integer, the number of power allocation patterns is finite, and we could traverse all possible patterns with a small number of users, e.g., $K = 2, 3$.
- **Numerical solution:** By calculate the optimization functions in Eq. (21) and (22) of all possible power allocation patterns, we could obtain the numerical solutions to P'_1 and P_2 .

For instance, by considering P'_1 with $K = 2$, the total power ρ is divided into $N = 200$ parts with each one $\frac{\rho}{N}$. By applying the constraints of T_k , we could obtain all $\frac{N}{2} - 1$ possible power allocation patterns, i.e., $T_1 = \frac{N}{2} + 1, \frac{N}{2} + 2, \dots, N$ and $T_2 = N - T_1$. The optimization function could be simplified as $f'_1(\rho_1) = f_{\text{SER}}\left(\frac{\rho_1}{\rho - \rho_1 + \sigma_n^2}\right) + f_{\text{SER}}\left(\frac{\rho - \rho_1}{\sigma_n^2}\right)$. Assuming that the numerical result f_{SER}^* is known (closed-form solution of f_{SER} is not available in literature [12]–[14]). Afterward, we calculate the optimization function $f'_1(\rho_1)$ for all $\frac{N}{2} - 1$ possible power allocation patterns, and the optimum one with the lowest value of f'_1 will be selected as the final power allocation result.

C. USER PAIRING

Different from the conventional NOMA to maximize the achievable rate of paired users regardless of the modulation issue [25]–[27], minimizing the SER of paired users is selected as the target in the proposed NOMA-SM. Hence, the results in conventional NOMA that users with more distinctive channel conditions are preferred to be paired could not be directly applied in the proposed NOMA-SM [27].

Due to the hardware cost of power amplifiers (PAs), fixed power allocation coefficients are usually deployed at the transmitter side [26]. In two user pairing case, we assume that the fixed power allocation coefficients are $\rho_1^* = \alpha\rho$, $\rho_2^* = (1 - \alpha)\rho$. By considering the near-far effect from user distribution, the channel matrix of the k -th user is

redefined as

$$\mathbf{G}^{(k)} = \sqrt{\beta_k} \mathbf{H}^{(k)}, \quad (24)$$

where β_k , $1 \leq k \leq K$ denote the large-scale fading coefficients of the k -th users [23]. Without loss of generality, we assume $\beta_1 \leq \dots \leq \beta_K$. Hence, for the k -th user with power allocation ρ_k , its equivalent power allocation could be regarded as $\beta_k \rho_k$. A simple yet efficient method is to pair users with equivalent power allocation approaching the optimization results obtained through the proposed low-complexity power allocation scheme. Specifically, for the k_1 -th user, the following principle could be applied to select its best pairing user as

$$k_2 = \arg \min_{k \neq k_1} \min \left\{ \left| \frac{\beta_{k_1} \rho_1^*}{\beta_k \rho_2^*} - \frac{\rho_1'}{\rho_2'} \right|, \left| \frac{\beta_k \rho_1^*}{\beta_{k_1} \rho_2^*} - \frac{\rho_1'}{\rho_2'} \right| \right\}, \quad (25)$$

where $\{k_1^*, k_2^*\}$ denotes the selected two pairing users among total K users and $\{\rho_1', \rho_2'\}$ denotes the numerical results provided by the proposed low-complexity power allocation scheme. Hence, with fixed power allocation, users with equivalent power allocation approaching the optimization results could be paired to achieve better SER performance balance in NOMA-SM.

V. SIMULATION RESULTS

In this section, Monte Carlo based simulation results are presented to compare the performance of OMA-SM, TAG-SM and the proposed NOMA-SM. For most cases, i.i.d. Rayleigh fading channel is considered, while the near-far effect is added in user pairing discussion for NOMA-SM. Firstly, the error propagation in NOMA-SM is presented to verify the SER analysis (Corollary 1). Afterwards, the proposed low-complexity power allocation is compared to the simulation results. Then, the BER performance of OMA-SM, TAG-SM and the proposed NOMA-SM with low-complexity power allocation is compared in detail. Finally, the performance of user pairing method based on Eq. (25) is verified.

A. ERROR PROPAGATION

Fig. 3 presents the error propagation by both theoretical analysis and numerical simulation with $K = 2, N_t = N_r = 4$, QPSK and i.i.d. Rayleigh fading channel. Since the closed-form solution of f_{SER} is not available in literature [12]–[14], the numerical result f_{SER}^* (a point-to-point QPSK based SM transmission with a 4×4 i.i.d. Rayleigh fading channel) in region $\frac{\rho}{\sigma_n^2} \in [-30, 30]$ dB is firstly acquired through numerical simulation. Specifically, the region $[-30, 30]$ dB is sampled by 600 points (sample interval is 0.1 dB) and the

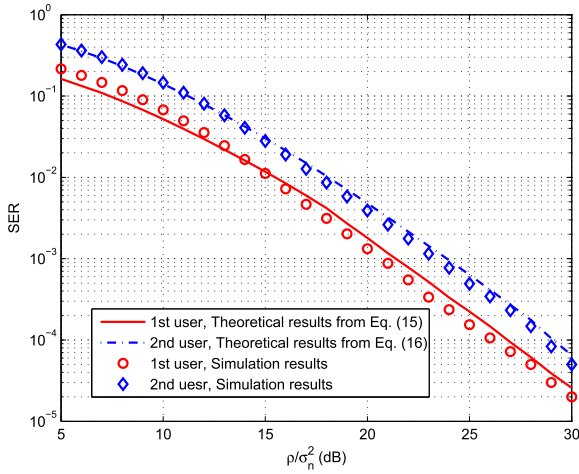


FIGURE 3. Error propagation among users in NOMA-SM with $K = 2$, $N_t = N_r = 4$, QPSK and $R_{\text{NOMA-SM},k} = 4$ bits/symbol.

SER performance among $[-30, 30]$ dB could be calculated by linear interpolation based on these 600 sampled points.

By assuming the same receiving SNR of the two users after SIC procedure, i.e., $\frac{\rho_1}{\rho_2 + \sigma_n^2} = \frac{\rho_2}{\sigma_n^2}$, the power allocation could be calculated as $\rho_2 = \sqrt{\sigma_n^4 + \sigma_n^2 \rho} - \sigma_n^2$, $\rho_1 = \rho - \rho_2$. With the numerical result f_{SER}^* , the theoretical SER performance of the 1st user and the 2nd user could be easily obtained through Eq. (15) and Eq. (16), which is presented in Fig. 3 by red and blue lines, respectively. Meanwhile, the SER performance of these two users by numerical simulation is also presented in Fig. 3 by red circle and blue diamond, respectively. The theoretical results anastomose with the simulation results well, which could be utilized to guide the design of power allocation and user pairing.

From Fig. 3, the SER performance of the 2nd user is always worse than the 1st user due to the error propagation in SIC procedure as shown in *Corollary 1*. Hence, similar to the conventional NOMA [23], [24], error propagation also leads performance degradation in the latter users in the proposed NOMA-SM, which indicates that the number of multiplexing users should be small, e.g., typically $K = 2$.

B. POWER ALLOCATION

Fig. 4 shows the power allocation results obtained through the proposed low-complexity scheme and the numerical simulation with $K = 2$, $N_t = N_r = 4$, QPSK and $R_{\text{NOMA-SM},k} = 4$ bits/symbol. Two different optimization targets are considered, i.e., P_1 to minimize the maximal SER of two users and P_2 to minimize the average SER. The total power ρ is uniformly divided into $N = 200$ parts, and the one with the lowest SER among all possible patterns will be selected as the final power allocation result. For numerical simulation method, the SER performance of each sample point has to be acquired based on a large number of channel realization (at least 10^6), which is drawn in Fig. 4 by red circle and blue diamond for P_1 and P_2 , respectively. Due to the unacceptable computational complexity, simulation based

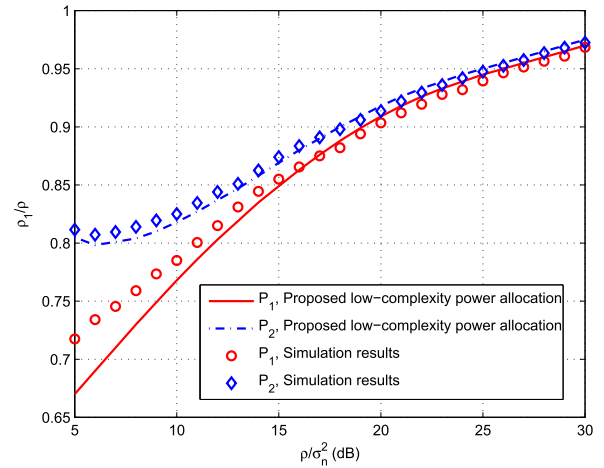


FIGURE 4. Power allocation for NOMA-SM with $K = 2$, $N_t = N_r = 4$, QPSK and $R_{\text{NOMA-SM},k} = 4$ bits/symbol.

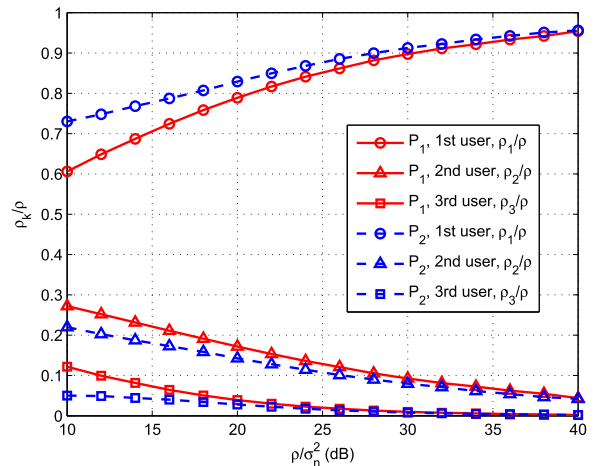


FIGURE 5. The proposed low-complexity power allocation result for NOMA-SM with $K = 3$, $N_t = N_r = 4$, QPSK and $R_{\text{NOMA-SM},k} = 4$ bits/symbol.

power allocation could not be deployed in practical NOMA-SM. While, in the proposed low-complexity power allocation scheme, based on the SER analysis and the numerical result f_{SER}^* , the target function could be directly calculated from Eq. (21) and Eq. (22) and the sample point with the lowest SER will be selected. Compared with simulation based power allocation, the computational complexity of the proposed scheme is almost negligible. From Fig. 4, the power allocation results of numerical simulation method and our proposal are similar and overlapped in high SNR region. Hence, the proposed low-complexity power allocation could be deployed to achieve the trade-off between computational complexity and system performance.

Fig. 5 presents the proposed low-complexity power allocation result for NOMA-SM with $K = 3$, $N_t = N_r = 4$, QPSK, and $R_{\text{NOMA-SM},k} = 4$ bits/symbol. For this $K = 3$ user multiplexing case, the total power is divided in a finer way with $N = 500$. All possible power allocation

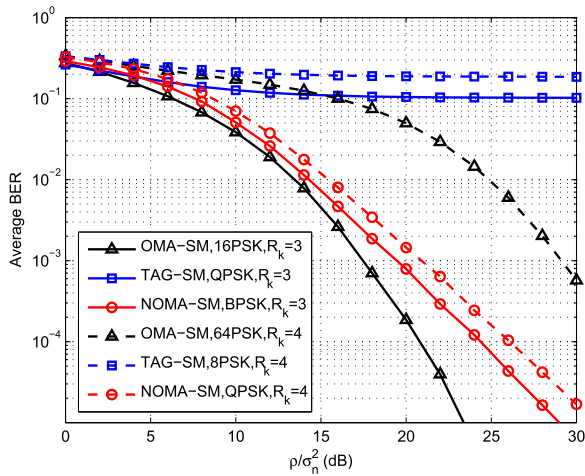


FIGURE 6. BER performance comparison of OMA-SM, TAG-SM and NOMA-SM against $\frac{\rho}{\sigma_n^2}$ with $K = 2$, $N_t = N_r = 4$ and $R_k = 3, 4$ bits/symbol.

patterns $\{T_1, T_2, T_3\}$ meeting the constraints $T_1 > T_2 > T_3$, $\sum_{k=1}^3 T_k = N$ will be traversed to calculate the optimization function, and the one with lowest value will be selected as the numerical results. As shown in Fig. 5, the low-complexity power allocation scheme could be also applied when the number of users is slightly increased, e.g., $K = 3$.

C. PERFORMANCE COMPARISON

Fig. 6 compares the average BER performance of OMA-SM, TAG-SM and the proposed NOMA-SM against the average receiving SNR $\frac{\rho}{\sigma_n^2}$ with $K = 2$, $N_t = N_r = 4$ and $R_k = 3, 4$ bits/symbol. For OMA-SM with $R_k = 3$ bits/symbol, $M = 2^{KR_k - \log_2(N_t)}$ (16PSK) has to be applied and the BER performance (black triangle line in Fig. 6) is better than TAG-SM and NOMA-SM (blue square line and red circle line in Fig. 6). However, by increasing R_k from 3 to 4 bits/symbol, much higher order PSK, i.e., 64PSK, has to be applied in OMA-SM, which leads to severe performance degradation (black triangle dashed line in Fig. 6). Hence, the most disadvantage of OMA-SM is the data rate limitation as shown in Eq. (2), which is caused by exclusive time-frequency resource allocation. For TAG-SM, the BER performance is acceptable at low SNR region, i.e., $\frac{\rho}{\sigma_n^2} < 5$ dB. By increasing the SNR $\frac{\rho}{\sigma_n^2}$, the BER performance of TAG-SM could not be improved due to the inter-user interference, which further leads to an error floor. According to [18]–[20], precoding and adaptive TA selection schemes could be adopted to improve the BER performance of TAG-SM, but at the cost of additional overhead to obtain CSI at the transmitter. For the proposed NOMA-SM, the low-complexity power allocation scheme is applied together. From Fig. 6 with $R_k = 3$ bits/symbol, similar BER performance of NOMA-SM with OMA-SM could be achieved. However, by increasing R_k from 3 to 4 bits/symbol, NOMA-SM only needs to replace BPSK by QPSK and the degradation of BER performance is quite little (two red circle lines in Fig. 6), which is much lower than both OMA-SM and TAG-SM.

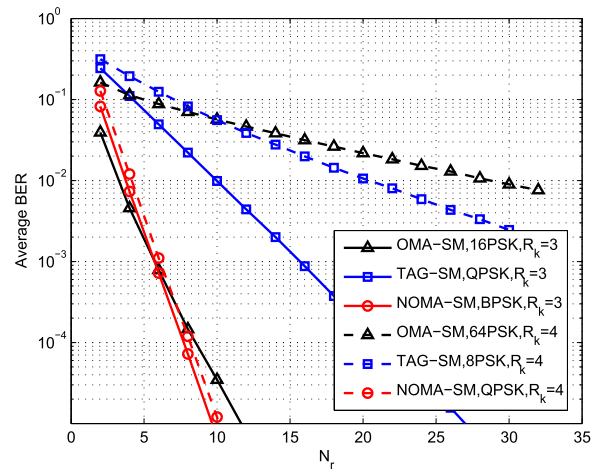


FIGURE 7. BER performance comparison of OMA-SM, TAG-SM and NOMA-SM against N_r with $K = 2$, $N_t = 4$ and $R_k = 3, 4$ bits/symbol.

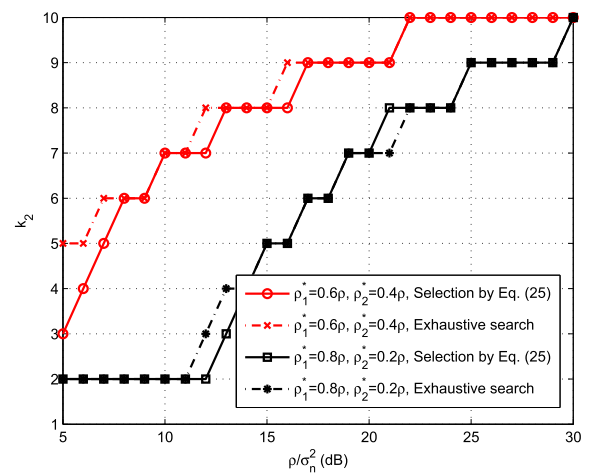


FIGURE 8. User pairing for NOMA-SM with user number $K = 10$.

Fig. 7 compares the average BER performance of OMA-SM, TAG-SM and the proposed NOMA-SM against the RA number N_r with $K = 2$, $N_t = 4$ and $R_k = 3, 4$ bits/symbol. By considering $R_k = 3$ bits/symbol, similar BER performance of OMA-SM and NOMA-SM could be observed ($N_r = 8$ to achieve $BER = 10^{-4}$) and could be significantly improved by increasing the RA number N_r . Meanwhile, a much larger number of RAs is required for TAG-SM to achieve the same BER performance, i.e., $N_r > 20$ to achieve $BER = 10^{-4}$. By increasing R_k to 4 bits/symbol, similar BER performance of the proposed NOMA-SM could be obtained as that with $R_k = 3$ bits/symbol. For TAG-SM, there is a BER performance degradation when the transmit rate R_k is increased from 3 to 4 bits/symbol, while a much larger BER performance degradation could be found in OMA-SM. These is no error floor of TAG-SM in Fig. 7 since the increasing of N_r could always improve the detection performance. In general, compared to both OMA-SM and TAG-SM, the proposed NOMA-SM is able to achieve better BER performance against the RA number N_r .

D. USER PAIRING

Fig. 8 presents the user pairing result in NOMA-SM with user number $K = 10$. According to Eq. (24), we firstly set the large-scale fading coefficients $[\beta_1 \beta_2 \cdots \beta_{10}] = [1 \ 0.9 \ 0.8 \ \cdots \ 0.1]$. For the 1st user, we aim to select its best pairing user with the lowest BER performance from the rest 9 users for NOMA-SM. Two cases of fixed power allocation are considered, i.e., $\rho_1^* = 0.6\rho$, $\rho_2^* = 0.4\rho$ and $\rho_1^* = 0.8\rho$, $\rho_2^* = 0.2\rho$. Based on exhaustive search among the rest 9 users, the best pairing user for the 1st user is plotted by dashed lines as shown in Fig. 8. Based on Eq. (25), the pairing user for the 1st user could be selected with low complexity to approach the power allocation result obtained from P_2 . The user pairing results based on Eq. (25) is plotted in Fig. 8 by solid lines, which is almost the same as the exhaustive search result.

VI. CONCLUSIONS

When SM is employed in DL-MU transmission scenarios, OMA-SM suffers from low spectral efficiency, while TAG-SM suffers from severe inter-user interference. In this paper, by cooperating NOMA with SM, NOMA-SM has been proposed to achieve a trade-off between spectral efficiency and interference mitigation. By deploying SIC detection at user side, inter-user interference could be effectively eliminated with the sacrifice of error propagation among users. SER approximation in NOMA-SM is derived and a low-complexity power allocation is proposed. Moreover, when near-far effect from user distribution is considered, user pairing for NOMA-SM is also addressed. The performance of OMA-SM, TAG-SM and the proposed NOMA-SM is compared by simulations and the effectiveness of the proposed low-complexity power allocation and user pairing is verified.

REFERENCES

- [1] J. G. Andrews et al., "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.
- [2] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
- [3] A. Fehske, G. Fettweis, J. Malmudin, and G. Biczok, "The global footprint of mobile communications: The ecological and economic perspective," *IEEE Commun. Mag.*, vol. 49, no. 8, pp. 55–62, Aug. 2011.
- [4] J. Liu, H. Guo, Z. M. Fadlullah, and N. Kato, "Energy consumption minimization for FiWi enhanced LTE-A HetNets with UE connection constraint," *IEEE Commun. Mag.*, vol. 54, no. 11, pp. 56–62, Nov. 2016.
- [5] P. Frenger, P. Moberg, J. Malmudin, Y. Jading, and I. Godor, "Reducing energy consumption in LTE with cell DTX," in *Proc. IEEE Veh. Technol. Conf. (VTC Spring)*, Yokohama, May 2011, pp. 1–5.
- [6] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-blast: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. Int. Symp. Signals, Syst. Electron. (ISSSE)*, Italy, Pisa, Oct. 1998, pp. 295–300.
- [7] H. Jafarkhani, *Space-Time Coding, Theory Practive*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [8] 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.
- [9] 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.
- [10] M. D. 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.
- [11] R. Mesleh, M. D. Renzo, H. Haas, and P. M. Grant, "Trellis coded spatial modulation," *IEEE Trans. Wireless Commun.*, vol. 9, no. 7, pp. 2349–2361, Jul. 2010.
- [12] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Spatial modulation: Optimal detection and performance analysis," *IEEE Commun. Lett.*, vol. 12, no. 8, pp. 545–547, Aug. 2008.
- [13] T. L. Narasimhan and A. Chockalingam, "On the capacity and performance of generalized spatial modulation," *IEEE Commun. Lett.*, vol. 20, no. 2, pp. 252–255, Feb. 2016.
- [14] M. D. 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.
- [15] T. L. 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.
- [16] S. Wang, Y. Li, and J. Wang, "Multiuser detection in massive spatial modulation MIMO with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2156–2168, Apr. 2015.
- [17] Z. Gao, L. Dai, Z. Wang, S. Chen, and L. Hanzo, "Compressive-sensing-based multiuser detector for the large-scale SM-MIMO uplink," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8725–8730, Oct. 2016.
- [18] X. Li, Y. Zhang, L. Xiao, X. Xu, and J. Wang, "A novel precoding scheme for downlink multi-user spatial modulation system," in *Proc. IEEE 24th Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, London, U.K., Sep. 2013, pp. 1361–1365.
- [19] S. Narayanan, M. J. Chaudhry, A. Stavridis, M. Di Renzo, F. Graziosi, and H. Haas, "Multi-user spatial modulation MIMO," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Turkey, Istanbul, Apr. 2014, pp. 671–676.
- [20] X. Wu, M. D. Renzo, and H. Haas, "A novel multiple access scheme based on spatial modulation MIMO," in *Proc. IEEE 19th Int. Workshop Comput. Aided Modeling Design Commun. Links Netw. (CAMAD)*, Greece, Athens, Dec. 2014, pp. 285–289.
- [21] X. Wu, M. D. Renzo, and H. Haas, "Adaptive selection of antennas for optimum transmission in spatial modulation," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3630–3641, Jul. 2015.
- [22] M. Maleki, H. R. Bahrami, and A. Alizadeh, "Layered spatial modulation for multiuser communications," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 7143–7159, Oct. 2016.
- [23] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. IEEE 77th Veh. Technol. Conf. (VTC Spring)*, Dresden, Germany, Jun. 2013, pp. 1–5.
- [24] A. Benjebbour, Y. Saito, Y. Kishiyama, A. Li, A. Harada, and T. Nakamura, "Concept and practical considerations of non-orthogonal multiple access (NOMA) for future radio access," in *Proc. IEEE Int. Symp. Intell. Signal Process. Commun. Syst. (ISPACS)*, Naha, Japan, Nov. 2013, pp. 770–774.
- [25] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, Dec. 2014.
- [26] S. Timotheou and I. Krikididis, "Fairness for non-orthogonal multiple access in 5G systems," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1647–1651, Oct. 2015.
- [27] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6010–6023, Aug. 2016.
- [28] J. Wang, S. Jia, and J. Song, "Signal vector based detection scheme for spatial modulation," *IEEE Commun. Lett.*, vol. 16, no. 1, pp. 19–21, Jan. 2012.
- [29] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*. New York, NY, USA: Springer, 2013.



XUDONG ZHU received the B.S. degree from the Department of Electronic Engineering, Tsinghua University, Beijing, China, in 2012, where he is currently pursuing the Ph.D. degree with the Broadband Communication and Signal Processing Laboratory. His main research interests include MIMO technique and mobile communication, especially in spatial modulation, sparse signal reconstruction, and massive MIMO.



ZHAOCHENG WANG (M'09–SM'11) received the B.S., M.S., and Ph.D. degrees from Tsinghua University, Beijing, China, in 1991, 1993, and 1996, respectively. From 1996 to 1997, he was a Post-Doctoral Fellow with Nanyang Technological University, Singapore. From 1997 to 1999, he was with OKI Techno Center, (Singapore) Pte. Ltd., Singapore, where he was a Research Engineer and later became a Senior Engineer. From 1999 to 2009, he was with Sony Deutschland GmbH, where he was a Senior Engineer and later became a Principal Engineer. He is currently a Professor of Electronic Engineering with Tsinghua University and also serves as the Director of Broadband Communication Key Laboratory, Tsinghua National Laboratory for Information Science and Technology. He has authored or co-authored over 120 journal papers. He holds 34 granted U.S./EU patents. He has co-authored two books, including *Millimeter Wave Communication Systems* (Wiley-IEEE Press), was selected by the IEEE Series on Digital and Mobile Communication. His research interests include wireless communications, visible light communications, millimeterwave communications, and digital broadcasting. He is a fellow of the Institution of Engineering and Technology. He served as an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2011 to 2015 and the IEEE Communications Letters from 2013 to 2016, and has also served as Technical Program Committee Co-Chair of various international conferences.



JIANFEI CAO received the B.S. degree from Hunan University in 2006, and the Ph.D. degree Beijing Jiaotong University, China, in 2006 and 2012, respectively. From 2009 to 2010, he was a Visiting Ph.D. student with the University of Southampton, U.K. From 2012 to 2014, he was an Associated Researcher with NEC Laboratory, China. From 2014 to 2016, he was a Senior Engineer with Samsung Research Laboratory China. Since 2016, he has been a Vice-Manager with the SONY China Research Laboratory. His research interests include FD-MIMO in LTE and massive MIMO in 5G NR.

• • •