

RESEARCH

Open Access



On the performance of receiver strategies for cooperative relaying cellular networks with NOMA

Jinjuan Ju^{1,2}, Guoan Zhang^{1*}, Qiang Sun^{1,3}, Li Jin¹ and Wei Duan¹

Abstract

In this paper, a receiver scheme for a cooperative multi-relay system with non-orthogonal multiple access (CMRS-NOMA) in cellular networks is proposed. In our proposed system, the base station (BS) would like to communicate with multiple users via multiple relays, and direct links between the BS and users are not considered. Specifically, we assume that the relays in the approaching range are allowed to assist the transmission since other relays are out of the cooperation range. In this way, each user and the corresponding relays can be implemented as one group. In addition, the source simultaneously transmits the superposition coded signals to all relay nodes. After receiving the signals, all the relay nodes decode the symbols by applying successive interference cancellation (SIC) and then reconstruct them into new NOMA signals, which will be forwarded to the users and decoded in a linear combination way. Note that the proposed reconstructions at the relay nodes are practical and simple, which leads to an advantageous decoding for the receivers. Moreover, the closed-form expressions in terms of the sum rate (SR) and the outage performance of the system are derived. Qualitative numerical results corroborating our theoretical analysis show that the proposed scheme significantly improves the performance in terms of the SR and outage probability compared with the existing works.

Keywords: Cooperative multi-relay system (CMRS), Non-orthogonal multiple access (NOMA), Sum rate (SR), Outage performance

1 Introduction

Recently, non-orthogonal multiple access (NOMA) has received considerable attention as a promising candidate for 5G systems, to realize an aggregate goal, such as superior spectral efficiency, balanced user fairness, intense connections, and low access latency [1, 2]. In contrast to the traditional orthogonal multiple access (OMA) scheme and recent studies [3–6], the NOMA technique allows multiple users to share time/frequency radio resources and allows distinguishing users by the different power levels [7–9]. In addition, a cooperative relaying system with the NOMA technique (i.e., CRS-NOMA) is proposed in [2], which not only improves spectral efficiency but also reduces the system complexity based on some special approaches. To further improve the system performance

in CRS-NOMA, the authors in [10] investigated a novel receiver design, where the performance in terms of the SR and outage probability are significantly improved. Since the distributions have an important impact, a new CRS-NOMA protocol is proposed in [11], in which the applications of simultaneous wireless information and power transfer (SWIPT) to NOMA networks are considered. Moreover, the works focusing on the power allocation schemes and different channel models, such as Rician fading channels and Nakagami- m fading channels in the NOMA technology, have also been studied in [12–14].

Due to the various advantages promised by NOMA technology in existing and future wireless communications systems, NOMA technology has been shown to be compatible with many other potential 5G techniques, such as massive multiple-input multiple-output (MIMO) and millimeter-wave (mmWave), as well as some applications in Internet of Things (IoT), such as vehicle-to-vehicle (V2V) communication scenarios and

*Correspondence: g Zhang@ntu.edu.cn

¹School of Information Science and Technology, Nantong University, Seyuan Road, Nantong, China

Full list of author information is available at the end of the article

machine-type communications (MTC). As an effective means to improve spectral efficiency, MIMO has been widely studied [15–18]. In [19–22], various algorithmic frameworks and performance optimization problems in MIMO-NOMA scenarios were also studied. However, most of the existing NOMA schemes rely on a key assumption that the channel conditions for the users are heterogeneous. To ensure that the potential of NOMA technology can be realized even if the users' channel conditions are similar, a MIMO-NOMA downlink transmission scenario has been considered, with one transmitter sending data to two users, e.g., an access point is serving two IoT devices [23]. Particularly, the authors considered that one user needs to be served quickly for small packet transmission, i.e., with a low targeted data rate, and the other user is to be served with the best effort. However, due to the scarcity of traditional wireless communication frequency resources, mmWave is slowly becoming popular because it has more bandwidth resources in the high-frequency band [24]. The authors in [25] investigated spatial and frequency wideband effects, called dual-wideband effects in massive MIMO systems, from the array signal processing perspective. By taking millimeter-wave-band communications into account, they described the transmission process to address the dual-wideband effects. In [26], to avoid the requirement that the base station (BS) should know all the channel state information (CSI) of the users, the applications of random beamforming to mmWave-NOMA systems were considered. In the application scenarios of IoT, such as V2V communications, since multiple antennas can be accommodated, large-scale MIMO becomes quite attractive [27, 28]. Inspired by the robustness of spatial modulation (SM) against channel correlation and the benefits of NOMA technology, the authors in [29] combined them into NOMA-SM to handle the effects of wireless V2V environments and to support the improvement in bandwidth efficiency.

The studies of NOMA technology and its compatibility with other technologies used in certain scenarios have clearly received considerable attention. For a scenario with multiple receiving nodes, both the transmission strategies and performance analysis are more practical and challenging for wireless communication systems. In addition, to the best of our knowledge, there are no relevant studies that focus on cellular networks with a multi-relay node transmission model with NOMA technology. Although the NOMA systems over Rician fading channels and/or Nakagami- m fading channels are more interesting and challenging, they are beyond the scope of this paper. Based on the above considerations, in this paper, a new transmission model named cooperative multi-relay system using NOMA (CMRS-NOMA) technology over Rayleigh fading channels is studied. To jointly decode the desired signals for each user, the new superposition coded signals are

reconstructed and broadcast by the multiple relay nodes. For assisting the multiple relay nodes in completing the transmission of the new superposition signals, a simple and practical decoding scheme with high efficiency and low complexity is investigated.

The implementations and main contributions of our proposed work are summarized as follows:

- A new CMRS-NOMA transmission scenario is proposed, in which the relays within the cooperation range are allowed to assist the transmissions from the BS to the users. Meanwhile, the direct transmission channels are not considered. After receiving signals, the relay nodes reconstruct the received signals into a new superposition code in a practical way, and the coded signal will be simply decoded with linear combinations at the receiver. Moreover, we also introduce a two-stage NOMA power allocation scheme at the BS and relay nodes, which is interesting and challenging.
- We investigate a new decoding method that can be described as a simple algorithm at each user after the relay nodes forward the reconstructed signals from the BS. Thanks to the reconstructions at the relay, each user can decode the desired signals with linear combinations, where the signals are forwarded from the corresponding relays. In addition, we qualitatively analyze a low-complexity, multi-stage power allocation strategy.
- The performance of the cooperative multi-relay NOMA technology in cellular networks is analyzed in terms of the ergodic SR and outage probability. Considering independent Rayleigh fading channels, we first derive the closed-form expressions of the ergodic SR and outage probability for a single-relay case. By extending the single-relay case into the remaining multi-relay node pair cases, a generalized reconstruction strategy is investigated, where all relay nodes follow the same regularity feature. For the remaining multi-relay node pair cases, in this manner, the power allocation factors of the proposed CMRS-NOMA should be changed in a circular manner, which is completely different from the traditional single-relay cooperative networks.
- Compared to the TDMA and CRS-NOMA schemes, the proposed CMRS-NOMA has better performances in terms of outage probability and spectral efficiency in both the analytical and numerical results. It is also observed that NOMA technology offers better fairness when multiple users' quality of services (QoS) are satisfied.

The remainder of this paper is organized as follows. In Section 2, the proposed new system model, power

allocation strategies, and decoding scheme of multi-relay cooperative NOMA for cellular networks are described. In Section 3, we analyze the performance in terms of the ergodic SR, outage probability, and some auxiliary indicators. In addition, closed-form expressions for the ergodic SR and outage performance are derived. In Section 4, we present and discuss the simulation results for our proposed works. This paper is concluded in Section 5.

2 System model and the proposed scheme

Consider a cooperative relaying communication system that consists of one source S ; multiple relay nodes, i.e., $R_1 \dots R_M$; and multiple users, $D_1 \dots D_N$, as shown in Fig. 1. We further consider that the BS wants to communicate with the users via the relay nodes since there are no direct links between them. In addition, the relays in the approaching range are allowed to assist the transmission, while the other relays are beyond the transmission range. It is assumed that all nodes are operating in a half-duplex mode and that each one is equipped with a single antenna. The channels between $S-R_i$ and R_i-D_j are denoted as h_{SR_i} and $h_{R_iD_j}$, respectively, which are assumed to be Rayleigh fading channels with variances β_{SR_i} and

$\beta_{R_iD_j}$. Furthermore, the CSI is assumed to be perfectly known to each receiving node.

There are two phases involved in our proposed scheme. In the first phase, the source broadcasts N symbols (s_1, s_2, \dots, s_N) simultaneously to the relays. By adopting the superposition code, the transmitted signal from the source is given as $\sum_{k=1}^N \sqrt{a_k P_t} s_k$, where s_k is the symbol for the k th ($1 \leq k \leq N$) user, a_k with $\sum_{k=1}^N a_k = 1$ denotes the power allocation coefficient, and P_t is the total transmit power. Similar to the traditional NOMA technology, the power coefficients are determined by the qualities of the $S-R_i$ channels. The observation at the i th ($1 \leq i \leq M, M = N$) relay in multiple relay nodes is given by

$$y_{R_i} = h_{SR_i} \sum_{k=1}^N \sqrt{a_k P_t} s_k + n_{R_i}, \tag{1}$$

where $n_{R_i} \sim \mathcal{CN}(0, \sigma_{R_i}^2)$ are the additive white Gaussian noises (AWGNs). Without loss of generality, following the NOMA decoding principle, $s_j (1 \leq j \leq N - 1)$ is decoded first and allocated with more transmit power, and then s_{j+1} is subsequently decoded. Successive detection will be performed at the i th relay at the end of this phase. Because

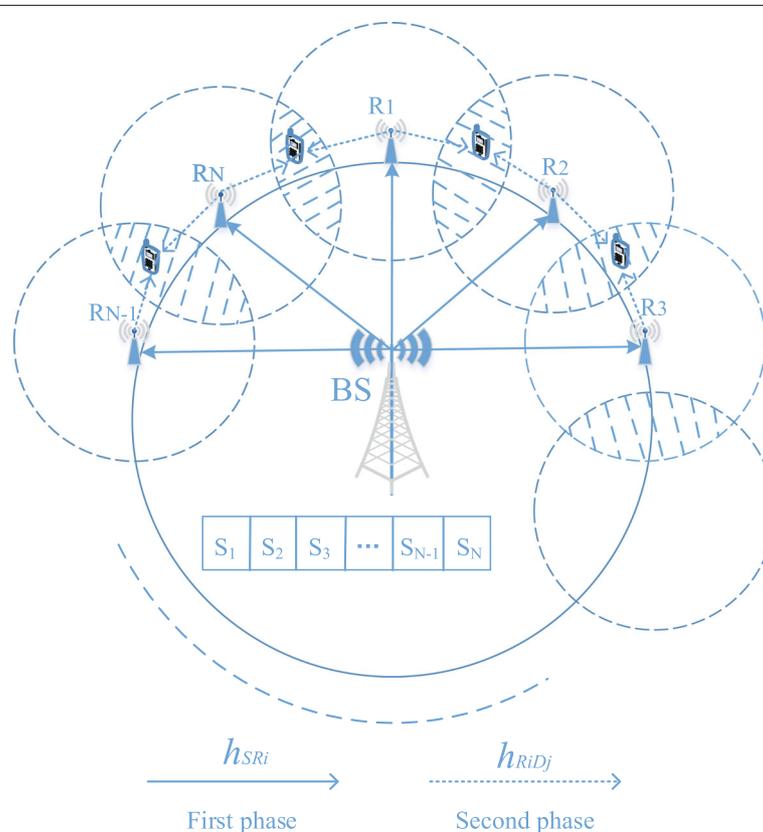


Fig. 1 The proposed multi-relay cooperative system with NOMA (MCRS-NOMA)

of the cellular distribution, all the S - R_i channels are spatially disparate and have distinct transmission qualities at this stage, i.e.,

$$h_{SR_1} \neq h_{SR_2} \neq \dots \neq h_{SR_N}, \quad (2)$$

In the second phase, with the superposition coding (SC) and successive interference cancellation (SIC), all the relay nodes are available to successively decode the symbols. In this manner, the received signal-to-interference-plus-noise ratio (SINR) at the i th relay node is given by

$$\gamma_{R_i}^{(s_k)} = \frac{|h_{SR_i}|^2 a_k P_t}{|h_{SR_i}|^2 \left(\sum_{m=k+1}^N a_m \right) P_t + \sigma^2}, \quad (3)$$

and

$$\gamma_{R_i}^{(s_N)} = \frac{|h_{SR_N}|^2 a_N P_t}{\sigma^2}. \quad (4)$$

To provide a simple and practical decoding, the relays will reconstruct and forward new superposition coded signals, and the user jointly decodes the information symbols from the received signals by employing the linear combination. In fact, only the signals that come from the approaching relays need to be provided for decoding the desired signal. Indeed, when transmitting the reconstructed signals, the power allocation strategy at the relay needs to be designed, which means that a power allocation is still required at each node. As shown in Fig. 1, there are N relay nodes, and the corresponding power allocation coefficients b_{R_i} with $\sum_{i=1}^N b_{R_i} = 1$ are considered in the system. At this time, considering the power reallocation at each relay node for transmission, third stage power allocation will occur with coefficients c_k with $\sum_{k=1}^N c_k = 1$, which will make the system analysis complex and not suitable for research. For simplicity, the information transmitted at each relay node will no longer be allocated in the relay layer but rather sent out with full power.

2.1 Single-relay scenario

To successfully analyze the multi-relay case, we first consider a single-relay case, which is a simple CRS consisting of only one source, one relay, and one destination. Assume that a direct link between the source and the destination exists, the destination jointly decodes the received signals from both the source and the relay.

At the first time slot, the BS simultaneously transmits the signal $\sqrt{a_1 P_t} s_1 + \sqrt{a_2 P_t} s_2$ to the relay and destination, where s_k denotes the broadcast symbol at the source and a_1 and a_2 with $a_1 + a_2 = 1$ are the power allocation factors. At the second time slot, a new symbol s_r with superposition coding $s_r = \sqrt{b_1 P_t} s_1 - \sqrt{b_2 P_t} s_2$ is forwarded from the relay node to the destination, where b_1 and b_2 with $b_1 + b_2 = 1$ are new power allocation coefficients. Note

that the subtraction is necessary, and it can perfectly cancel the interference signals at the destination. The received signals at the destination can be written as

$$y_D^{(1)} = h_{SD} \left(\sqrt{a_1 P_t} s_1 + \sqrt{a_2 P_t} s_2 \right) + n_D^{(1)}, \quad (5)$$

$$y_D^{(2)} = h_{RD} \left(\sqrt{b_1 P_t} s_1 - \sqrt{b_2 P_t} s_2 \right) + n_D^{(2)}, \quad (6)$$

where $n_D^{(i)}$ is the AWGN at the destination during the i th time slot with zero mean and variance σ^2 . To jointly decode s_1 and s_2 at the destination, a linear combination is employed to produce interference-free signals $y_D^{(1)} \sqrt{b_2} h_{RD} + y_D^{(2)} \sqrt{a_2} h_{SD}$ and $y_D^{(1)} \sqrt{b_1} h_{RD} - y_D^{(2)} \sqrt{a_1} h_{SD}$. Therefore, the target signals can be expressed as

$$T_{s_1} = h_{SD} h_{RD} \zeta \sqrt{P_t} s_1 + \sqrt{b_2} h_{RD} n_D^{(1)} + \sqrt{a_2} h_{SD} n_D^{(2)}, \quad (7)$$

and

$$T_{s_2} = h_{SD} h_{RD} \zeta \sqrt{P_t} s_2 + \sqrt{b_1} h_{RD} n_D^{(1)} - \sqrt{a_1} h_{SD} n_D^{(2)}, \quad (8)$$

where $\zeta = \sqrt{a_1 b_2} + \sqrt{a_2 b_1}$. With (7) and (8), the corresponding effective SNRs for s_1 and s_2 are given by

$$\gamma_D^{(s_1)} = \frac{|h_{SD}|^2 |h_{RD}|^2 \zeta^2 \rho}{b_2 |h_{RD}|^2 + a_2 |h_{SR}|^2}, \quad (9)$$

and

$$\gamma_D^{(s_2)} = \frac{|h_{SD}|^2 |h_{RD}|^2 \zeta^2 \rho}{b_1 |h_{RD}|^2 + a_1 |h_{SR}|^2}, \quad (10)$$

respectively, where $\rho = \frac{P_t}{\sigma^2}$ denotes the transmit SNR.

2.2 Proposed multi-relay scheme

In this subsection, we extend the single-relay case into the case with multiple users and relays, i.e., CMRS-NOMA. Note that direct links between the BS and the users are not considered since such consideration will lead to extreme complexity for the power allocation. Moreover, in this study, our aim is to demonstrate that our proposed reconstruction at the relay node has advantages compared with the directly forwarding one.

In this scheme, in the first phase, the source broadcasts N symbols $\sum_{k=1}^N \sqrt{a_k P_t} s_k$ only to the relay nodes and the received signal at the i th relay in multiple relay nodes, as shown in Eq. (1). By employing SIC, all the relay nodes can decode the symbols successively. In the second phase, the relay nodes reconstruct new superposition coded signals by using linear operations. In contrast to the single-relay case, there are multiple relay nodes, where different relay nodes will reconstruct in a different way by using different linear operations and forward them to the user in a short-range circular communication area. In this process, a two-stage NOMA power allocation scheme at the BS and the relay nodes is introduced. The choice of decoding method at each user after the relay nodes forward

new NOMA signals directly affects whether the interference signals are perfectly cancelled. Because the relays in the approaching range are allowed to assist the transmission, each user and the corresponding relay nodes can be implemented as one group. By using simple addition and division, all the other signals except the desired one will be removed.

Our proposed linear operations at the relay nodes can be described as a simple algorithm. The detailed steps of the linear algorithm are as follows:

- **Step 1:** We first choose a relay in the direction of 90° of the coordinate axis and mark it as the number 1 relay; then, along the clockwise direction, we mark the remaining relays as 2,3... until the last relay is marked as N . Assume that multiple relay nodes are distributed in a cellular around the BS with equal distances from the relays to the BS and that the last relay node is connected with the first one.
- **Step 2:** To handle the operational demands in the distributed relay nodes, we place all the adjacent relay node pairs into a node pair list (NPL) and identify the corresponding users within the overlapping area of the relay node pairs at the destinations that are outside of the BS service area. The NPL and CU of step 2 are given in Table 1.
- **Step 3:** For the node pairs that have been served, we obtain each new symbol $s_{R_i} (1 \leq i \leq N)$ from a simple linear combination in sequence and forward them to the area with the full power of a single antenna. The reconstructed signal combinations for each relay are shown in the following group.

$$\begin{cases} s_{R_1} = \sqrt{b_{R_1} a_1 P_t} s_1 + \sqrt{b_{R_1} a_2 P_t} s_2 + \sqrt{b_{R_1} a_3 P_t} s_3 + \dots + \sqrt{b_{R_1} a_N P_t} s_N & (a) \\ s_{R_2} = \sqrt{b_{R_2} a_1 P_t} s_1 - \sqrt{b_{R_2} a_2 P_t} s_2 - \sqrt{b_{R_2} a_3 P_t} s_3 - \dots - \sqrt{b_{R_2} a_N P_t} s_N & (b) \\ s_{R_3} = \sqrt{b_{R_3} a_1 P_t} s_1 + \sqrt{b_{R_3} a_2 P_t} s_2 - \sqrt{b_{R_3} a_3 P_t} s_3 - \dots - \sqrt{b_{R_3} a_N P_t} s_N & (c) \\ s_{R_4} = \sqrt{b_{R_4} a_1 P_t} s_1 + \sqrt{b_{R_4} a_2 P_t} s_2 + \sqrt{b_{R_4} a_3 P_t} s_3 - \dots - \sqrt{b_{R_4} a_N P_t} s_N & (d) \\ \vdots & \\ s_{R_N} = \sqrt{b_{R_N} a_1 P_t} s_1 + \sqrt{b_{R_N} a_2 P_t} s_2 + \sqrt{b_{R_N} a_3 P_t} s_3 + \dots - \sqrt{b_{R_N} a_N P_t} s_N & (e) \end{cases} \quad (11)$$

- **Step 4:** After we finish reconstructing the new signals for each relay and send them out, we can jointly decode the received signals from two adjacent relays

Table 1 RNP and corresponding user

Relay node pair	Corresponding user
(R_1, R_2)	$(D_1, user_1)$
(R_2, R_3)	$(D_2, user_2)$
...	...
(R_{N-1}, R_N)	$(D_{N-1}, user_{N-1})$
(R_N, R_1)	$(D_N, user_N)$

by employing a simple linear combination method at the users. As shown in the above expressions, for example, signal 1 is obtained by adding $(s_{R_1} \cdot \sqrt{b_{R_2}})$ to $(s_{R_2} \cdot \sqrt{b_{R_1}})$ and then dividing by $(2 \cdot \sqrt{b_{R_1} b_{R_2} a_1 P_t})$. Signal 2 is achieved by subtracting $(s_{R_2} \cdot \sqrt{b_{R_3}})$ from $(s_{R_3} \cdot \sqrt{b_{R_2}})$ and then dividing by $(2 \cdot \sqrt{b_{R_2} b_{R_3} a_2 P_t})$. The remainder can be achieved in the same manner, and the last signal N will be acquired by also deducting the end equation, i.e., $(s_{R_N} \cdot \sqrt{b_{R_1}})$, from the first, i.e., $(s_{R_1} \cdot \sqrt{b_{R_N}})$, and then dividing by $(2 \cdot \sqrt{b_{R_1} b_{R_N} a_N P_t})$. Afterwords, all the signals will have been successfully decoded.

Taking signal 1 as an example, we analyze the received signal of user₁ that is cooperated by the number 1 and number 2 relays. Because a_k and b_{R_i} in group (11) are power allocation factors that meet their respective conditions, i.e., $\sum_{k=1}^N a_k = 1$ and $\sum_{i=1}^N b_{R_i} = 1$, group (11) can be simplified by allowing $c_{ik} = b_{R_i} \times a_k$, and formulae (11-a) and (11-b) can be described as

$$s_{R_1} = \sqrt{c_{11} P_t} s_1 + \sqrt{c_{12} P_t} s_2 + \dots + \sqrt{c_{1(N-1)} P_t} s_{N-1} + \sqrt{c_{1N} P_t} s_N, \quad (12)$$

and

$$s_{R_2} = \sqrt{c_{21} P_t} s_1 - \sqrt{c_{22} P_t} s_2 - \dots - \sqrt{c_{2(N-1)} P_t} s_{N-1} - \sqrt{c_{2N} P_t} s_N. \quad (13)$$

In our proposed system, the relay nodes forward new reconstructed symbols s_{R_i} with superposition signals to users as shown in the above formulae. Note that signals from two adjacent relay nodes only need to be processed simply to obtain the desired signal. With (12) and (13), the received signal at user₁ can be written as

$$\begin{aligned} y_{D_1}^{(R_1)} &= h_{R_1 D_1} \times s_{R_1} + n_{D_1}^{(R_1)} \\ &= h_{R_1 D_1} \times \left(\sqrt{c_{11} P_t} s_1 + \sum_{k=2}^N \sqrt{c_{1k} P_t} s_k \right) + n_{D_1}^{(R_1)}, \quad (14) \end{aligned}$$

$$\begin{aligned} y_{D_1}^{(R_2)} &= h_{R_2 D_1} \times s_{R_2} + n_{D_1}^{(R_2)} \\ &= h_{R_2 D_1} \times \left(\sqrt{c_{21} P_t} s_1 - \sum_{k=2}^N \sqrt{c_{2k} P_t} s_k \right) + n_{D_1}^{(R_2)}, \quad (15) \end{aligned}$$

where $n_{D_1}^{(R_i)}$ is the AWGN at user₁ from the i th relay node during the second time slot with zero mean and variance σ^2 . From the above, we find that there are two signals received at user₁, which are $y_{D_1}^{(R_1)}$ and $y_{D_1}^{(R_2)}$. As the single-relay model scheme, user₁ jointly decodes s_1 . With a simple linear combination, we could produce

interference-free signal $y_{D_1}^{(R_1)} \sqrt{b_{R_2}} h_{R_2 D_1} + y_{D_1}^{(R_2)} \sqrt{b_{R_1}} h_{R_1 D_1}$ for s_1 at user₁, with the target signal T_{s_1} as

$$T_{s_1} = h_{R_1 D_1} h_{R_2 D_1} \varsigma \sqrt{P_t s_1} + \sqrt{b_{R_2}} h_{R_2 D_1} n_{D_1}^{(R_1)} + \sqrt{b_{R_1}} h_{R_1 D_1} n_{D_1}^{(R_2)}, \quad (16)$$

where $\varsigma = \sqrt{c_{11} b_{R_2}} + \sqrt{c_{21} b_{R_1}} = 2\sqrt{b_{R_1} b_{R_2} a_1}$. If we perform symbol-by-symbol detection based on the output signals (16), the corresponding effective SNR for s_1 is given by

$$\gamma_{D_1}^{(s_1)} = \frac{|h_{R_1 D_1}|^2 |h_{R_2 D_1}|^2 \varsigma^2 \rho}{b_{R_2} |h_{R_2 D_1}|^2 + b_{R_1} |h_{R_1 D_1}|^2}, \quad (17)$$

where $\rho = \frac{P_t}{\sigma^2}$ denotes the transmit SNR.

Clearly, signal 2 to signal $N - 1$ can be obtained in a similar manner, i.e., for s_k ($2 \leq k \leq N - 1$) at user i , the interference-free signals $y_{D_i}^{(R_{i+1})} \sqrt{b_{R_i}} h_{R_i D_i} - y_{D_i}^{(R_i)} \sqrt{b_{R_{i+1}}} h_{R_{i+1} D_i}$ could also be generated by employing the same method. As described in step 4 above, the interference-free signal for s_N at user N (the last signal N) will be generated by $y_{D_N}^{(R_1)} \sqrt{b_{R_N}} h_{R_N D_N} - y_{D_N}^{(R_N)} \sqrt{b_{R_1}} h_{R_1 D_N}$. Similar to the previous descriptions, the target signal T_{s_i} ($2 \leq i \leq N - 1$) can be expressed as

$$T_{s_i} = h_{R_i D_i} h_{R_{i+1} D_i} \varsigma \sqrt{P_t s_i} + \sqrt{b_{R_{i+1}}} h_{R_{i+1} D_i} n_{D_i}^{(R_i)} + \sqrt{b_{R_i}} h_{R_i D_i} n_{D_i}^{(R_{i+1})}, \quad (18)$$

where $\varsigma = 2\sqrt{b_{R_i} b_{R_{i+1}} a_i}$; hence, the expressions of interference-free signal for T_{s_i} ($1 \leq i \leq N - 1$) are the same. However, the last one is slightly different from the ones in front; it is

$$T_{s_N} = h_{R_1 D_N} h_{R_N D_N} \varsigma \sqrt{P_t s_N} + \sqrt{b_{R_N}} h_{R_N D_N} n_{D_N}^{(R_1)} + \sqrt{b_{R_1}} h_{R_1 D_N} n_{D_N}^{(R_N)}, \quad (19)$$

where $\varsigma = 2\sqrt{b_{R_N} b_{R_1} a_N}$. Therefore, the corresponding effective SNRs for s_i ($1 \leq i \leq N - 1$) and s_N are given as

$$\gamma_{D_i}^{(s_i)} = \frac{4 |h_{R_{i+1} D_i}|^2 |h_{R_i D_i}|^2 (b_{R_i} b_{R_{i+1}} a_i) \rho}{b_{R_{i+1}} |h_{R_{i+1} D_i}|^2 + b_{R_i} |h_{R_i D_i}|^2}, \quad (20)$$

and

$$\gamma_{D_N}^{(s_N)} = \frac{4 |h_{R_1 D_N}|^2 |h_{R_N D_N}|^2 (b_{R_N} b_{R_1} a_N) \rho}{b_{R_N} |h_{R_N D_N}|^2 + b_{R_1} |h_{R_1 D_N}|^2}, \quad (21)$$

where $\rho = \frac{P_t}{\sigma^2}$ denotes the transmit SNR.

3 Performance analysis

In this section, we will analyze the performance of the achievable ergodic SR and the outage probability to characterize the superiority of our proposed scheme.

3.1 Ergodic SR analysis

According to Shannon's theorem, the achievable rate for symbol s_1 can be obtained from $\gamma_{R_1}^{(s_1)}$, $\gamma_{R_2}^{(s_1)}$, and $\gamma_{D_1}^{(s_1)}$ as

$$C_{s_1} = \frac{1}{2} \log_2 \left(1 + \min \left\{ \gamma_{R_1}^{(s_1)}, \gamma_{R_2}^{(s_1)}, \gamma_{D_1}^{(s_1)} \right\} \right), \quad (22)$$

where $\frac{1}{2}$ results from the two-hop transmission. Without loss of generality, the received signals corresponding to the relay node pairs can be described as

$$C_{s_i} = \frac{1}{2} \log_2 \left(1 + \min \left\{ \gamma_{R_i}^{(s_i)}, \gamma_{R_{i+1}}^{(s_i)}, \gamma_{D_i}^{(s_i)} \right\} \right). \quad (23)$$

Therefore, the achievable SR can be expressed as

$$C_{sum} = \sum_{i=1}^N C_{s_i}. \quad (24)$$

Denoting $|h_{SR_i}|^2 = v_{SR_i}$, $|h_{R_i D_j}|^2 = v_{R_i D_j}$, $d_i = b_{R_i} b_{R_{i+1}} a_i$, and $\tau = \sum_{m=i+1}^N a_m$, we have

$$\begin{aligned} & \min \left\{ \gamma_{R_i}^{(s_i)}, \gamma_{R_{i+1}}^{(s_i)}, \gamma_{D_i}^{(s_i)} \right\} \\ &= \min \left\{ \frac{|h_{SR_i}|^2 a_i \rho}{|h_{SR_i}|^2 \tau \rho + 1}, \frac{|h_{SR_{i+1}}|^2 a_i \rho}{|h_{SR_{i+1}}|^2 \tau \rho + 1}, \frac{4 |h_{R_i D_i}|^2 |h_{R_{i+1} D_i}|^2 d_i \rho}{b_{R_{i+1}} |h_{R_{i+1} D_i}|^2 + b_{R_i} |h_{R_i D_i}|^2} \right\} \\ &= \min \left\{ \frac{v_{SR_i} a_i \rho}{v_{SR_i} \tau \rho + 1}, \frac{v_{SR_{i+1}} a_i \rho}{v_{SR_{i+1}} \tau \rho + 1}, \frac{4 v_{R_i D_i} v_{R_{i+1} D_i} d_i \rho}{b_{R_{i+1}} v_{R_{i+1} D_i} + b_{R_i} v_{R_i D_i}} \right\}, \end{aligned} \quad (25)$$

where $i \in [1, N - 1]$ and $i \neq m$. For the case of $i = N$, the achievable rate C_{s_N} can be written as

$$C_{s_N} = \frac{1}{2} \log_2 \left(1 + \min \left\{ \gamma_{R_N}^{(s_N)}, \gamma_{R_1}^{(s_N)}, \gamma_{D_N}^{(s_N)} \right\} \right). \quad (26)$$

According to (21), $d_N = b_{R_N} b_{R_1} a_N$, for $i = N$, we have

$$\begin{aligned} & \min \left\{ \gamma_{R_N}^{(s_N)}, \gamma_{R_1}^{(s_N)}, \gamma_{D_N}^{(s_N)} \right\} \\ &= \min \left\{ |h_{SR_N}|^2 a_N \rho, |h_{SR_1}|^2 a_N \rho, \frac{4 |h_{R_N D_N}|^2 |h_{R_1 D_N}|^2 d_N \rho}{b_{R_1} |h_{R_1 D_N}|^2 + b_{R_N} |h_{R_N D_N}|^2} \right\} \\ &= \min \left\{ v_{SR_N} a_N \rho, v_{SR_1} a_N \rho, \frac{4 v_{R_N D_N} v_{R_1 D_N} d_N \rho}{b_{R_1} v_{R_1 D_N} + b_{R_N} v_{R_N D_N}} \right\}. \end{aligned} \quad (27)$$

Letting $X = \min \left\{ \gamma_{R_i}^{(s_i)}, \gamma_{R_{i+1}}^{(s_i)}, \gamma_{D_i}^{(s_i)} \right\}$, from (25) in which $i \neq N$, the complementary cumulative distribution function (CCDF) of X can be formulated as

$$\bar{F}_X(x) = Pr \left\{ \frac{v_{SR_i} a_i \rho}{v_{SR_i} \tau \rho + 1} > x, \frac{v_{SR_{i+1}} a_i \rho}{v_{SR_{i+1}} \tau \rho + 1} > x, \frac{4 v_{R_i D_i} v_{R_{i+1} D_i} d_i \rho}{b_{R_{i+1}} v_{R_{i+1} D_i} + b_{R_i} v_{R_i D_i}} > x \right\}. \quad (28)$$

Note that the CCDF of $\bar{F}_{v(\delta)}$ is $e^{-\frac{x}{\beta \delta}}$, for $\delta \in \{SR_i, R_i D_j\}$, Eq. (28) is equivalent to (29), as shown at the top of the

next page, where $Pr\{A|B\}$ denotes the conditional probability of event A on event B, and $E[\cdot]$ stands for statistical expectation.

$$\begin{aligned} \bar{F}_X(x) &= \bar{F}_{SR_i} \left(\frac{x}{a_i \rho - \tau \rho x} \right) \bar{F}_{SR_{i+1}} \left(\frac{x}{a_i \rho - \tau \rho x} \right) \\ &\quad \left(E_{v_{R_{i+1}D_i}} \left[Pr \left\{ v_{R_i D_i} > \frac{b_{R_{i+1}} v_{R_{i+1}D_i} x}{4 v_{R_{i+1}D_i} d_i \rho - b_{R_i} x} \mid v_{R_{i+1}D_i} > \frac{b_{R_i} x}{4 d_i \rho} \right\} \right] \right. \\ &\quad \left. + E_{v_{R_{i+1}D_i}} \left[Pr \left\{ v_{R_i D_i} < \frac{b_{R_{i+1}} v_{R_{i+1}D_i} x}{4 v_{R_{i+1}D_i} d_i \rho - b_{R_i} x} \mid v_{R_{i+1}D_i} < \frac{b_{R_i} x}{4 d_i \rho} \right\} \right] \right) \\ &= \left\{ \frac{1}{\beta_{R_{i+1}D_i}} e^{-\frac{x}{|a_i \rho - \tau \rho x| \beta_{SR_i}} - \frac{x}{|a_i \rho - \tau \rho x| \beta_{SR_{i+1}}} - \frac{b_{R_i} x}{4 d_i \rho \beta_{R_{i+1}D_i}}} \right. \\ &\quad \left. \int_{\frac{b_{R_i} x}{4 d_i \rho}}^{\infty} e^{-\frac{b_{R_{i+1}} u x}{(4 d_i \rho u - b_{R_i} x) \beta_{R_i D_i}} - \frac{u}{\beta_{R_{i+1}D_i}}} du \right\} \\ &\quad + \left\{ e^{-\frac{x}{|a_i \rho - \tau \rho x| \beta_{SR_i}} - \frac{x}{|a_i \rho - \tau \rho x| \beta_{SR_{i+1}}}} \left(1 - e^{-\frac{b_{R_i} x}{4 d_i \rho \beta_{R_{i+1}D_i}}} \right) \right. \\ &\quad \left. \int_0^{\frac{b_{R_i} x}{4 d_i \rho}} 1 - \frac{1}{\beta_{R_{i+1}D_i}} e^{-\frac{b_{R_{i+1}} u x}{(4 d_i \rho u - b_{R_i} x) \beta_{R_i D_i}} - \frac{u}{\beta_{R_{i+1}D_i}}} du \right\}, \end{aligned} \tag{29}$$

Since Eq. (29) is too complex to calculate, the high transmit SNR case is considered to simplify the calculation, i.e., $\rho \gg 1$, and we have the following approximations

$$\frac{v_{SR_i} a_i \rho}{v_{SR_i} \tau \rho + 1} \sim \frac{a_i}{\tau}, \tag{30}$$

and

$$\frac{v_{SR_{i+1}} a_i \rho}{v_{SR_{i+1}} \tau \rho + 1} \sim \frac{a_i}{\tau}. \tag{31}$$

Letting $t = 4d_i \rho u - b_{R_i} x$ and using $\int_0^\infty e^{-\frac{A}{x} - Bx} dx = 2\sqrt{\frac{A}{B}} K_1(2\sqrt{AB})$ [30]—Eq. (3.324.1), when $x < \frac{a_i}{\tau}$ ($\tau = \sum_{m=i+1}^N a_m$), (29) can be equivalently written as

$$\begin{aligned} \bar{F}_X(x) &= \frac{e^{-\frac{x}{4d_i \rho} \left(\frac{b_{R_{i+1}}}{\beta_{R_i D_i}} + \frac{b_{R_i}}{\beta_{R_{i+1}D_i}} \right)}}{4\beta_{R_{i+1}D_i} d_i \rho} \int_0^\infty e^{-\frac{b_{R_i} b_{R_{i+1}} x^2}{4d_i \rho \beta_{R_i D_i}} - \frac{t}{4d_i \rho \beta_{R_{i+1}D_i}}} dt \\ &= \frac{x}{\xi} e^{-x\phi} K_1 \left(\frac{x}{\xi} \right), \end{aligned} \tag{32}$$

where $\xi = 2d_i \rho \sqrt{\frac{\beta_{R_i D_i} \cdot \beta_{R_{i+1}D_i}}{b_{R_i} \cdot b_{R_{i+1}}}}$, $\phi = \frac{1}{4d_i \rho} \left(\frac{b_{R_{i+1}}}{\beta_{R_i D_i}} + \frac{b_{R_i}}{\beta_{R_{i+1}D_i}} \right)$, and $K_1(\cdot)$ denotes the first-order modified Bessel function of the second kind [31]. For the case $x > \frac{a_i}{\tau}$

($\tau = \sum_{m=i+1}^N a_m$), $\bar{F}_X(x) = 0$ always holds due to Eqs. (30) and (31). With (32), the achievable ergodic rate can be calculated as (33).

$$\begin{aligned} \tilde{C}_{S_i} &= \int_0^{\frac{a_i}{\tau}} \frac{1}{2} \log_2(1+x) dF_X(x) + \frac{1}{2} \log_2 \left(1 + \frac{a_i}{\tau} \right) \left(1 - F_X \left(\frac{a_i}{\tau} \right) \right) \\ &= \frac{1}{2} \log_2 \left(1 + \frac{a_i}{\tau} \right) - \frac{1}{2 \ln 2} \int_0^{\frac{a_i}{\tau}} \frac{1}{1+x} \left(1 - \frac{x}{\xi} e^{-x\phi} K_1 \left(\frac{x}{\xi} \right) \right) dx \\ &= \int_0^{\frac{a_i}{\tau}} \frac{1}{2 \ln 2(1+x)} \frac{x}{\xi} e^{-x\phi} K_1 \left(\frac{x}{\xi} \right) dx. \end{aligned} \tag{33}$$

Using the equality of $\int_0^\infty \frac{1}{2} \log_2(1+x) f_X(x) dx = \frac{1}{2 \ln 2} \int_0^\infty \frac{1 - F_X(x)}{1+x} dx$ and $F_X(x) = 1 - \bar{F}_X(x)$ and the approximation $K_\nu(x) \approx \frac{\Gamma(\nu)}{2} \left(\frac{2}{x} \right)^\nu$ [32]—Eq. (9.69) in the case of small x , where $\Gamma(\cdot)$ denotes the gamma function, (33) can be approximately rewritten as

$$\begin{aligned} \tilde{C}_{S_i} &\sim \frac{1}{2 \ln 2} \int_0^{\frac{a_i}{\tau}} \frac{1}{1+x} e^{-x\phi} dx \\ &= \frac{e^\phi}{2 \ln 2} \left[Ei \left(-\phi \frac{a_i}{\tau} - \phi \right) - Ei \left(-\phi \right) \right], \end{aligned} \tag{34}$$

where $\int_0^\infty \frac{e^{-\mu x} dx}{x + \psi} = e^{\mu \psi} [Ei(\mu \psi) - Ei(-\mu \psi)]$ [31]—Eq. (3.352.1) is used, and $Ei(\cdot)$ denotes the exponential integral function. The approximate closed expression obtained from the above analysis only satisfies the case where $i \in [1, N - 1]$.

By using a similar analysis method, letting $Y = \min \{ \gamma_{R_N}^{(SN)}, \gamma_{R_1}^{(SN)}, \gamma_{D_N}^{(SN)} \}$, from (27) in which $i = N$, the CCDF of Y can be formulated as

$$\begin{aligned} \bar{F}_Y(y) &= Pr \left\{ v_{SR_N} a_N \rho > y, v_{SR_1} a_N \rho > y, \frac{4v_{R_N D_N} v_{R_1 D_N} d_N \rho}{b_{R_1} v_{R_1 D_N} + b_{R_N} v_{R_N D_N}} > y \right\} \\ &= \bar{F}_Y^{(1)}(y) \times \bar{F}_Y^{(2)}(y) \times \bar{F}_Y^{(3)}(y) \\ &= e^{-\frac{y}{a_N \beta_{SR_N} \rho}} \times e^{-\frac{y}{a_N \beta_{SR_1} \rho}} \times \bar{F}_Y^{(3)}(y), \end{aligned} \tag{35}$$

where $\bar{F}_Y^{(1)}(y) = Pr \{ v_{SR_N} a_N \rho > y \}$, $\bar{F}_Y^{(2)}(y) = Pr \{ v_{SR_1} a_N \rho > y \}$ and $\bar{F}_Y^{(3)}(y) = Pr \left\{ \frac{4v_{R_N D_N} v_{R_1 D_N} d_N \rho}{b_{R_1} v_{R_1 D_N} + b_{R_N} v_{R_N D_N}} > y \right\}$. Letting $\lambda_1 = \frac{2b_{R_N}}{4d_N \rho \beta_{R_1 D_N}}$ and $\lambda_2 = \frac{2b_{R_1}}{4d_N \rho \beta_{R_N D_N}}$, since $\bar{F}_Y^{(3)}(y)$ could be written as $\bar{F}_Y^{(3)}(y) = Pr \left\{ \frac{2d_N \rho}{b_{R_N} b_{R_1}} \times \frac{2(b_{R_N} v_{R_N D_N})(b_{R_1} v_{R_1 D_N})}{b_{R_N} v_{R_N D_N} + b_{R_1} v_{R_1 D_N}} > y \right\}$, with the help of the cumulative distribution function(CDF) of the harmonic mean of two exponential variables, i.e., assuming that Z_1 and Z_2 are two independent exponential variables with parameters λ_1 and λ_2 respectively, the CDF of $Z = \frac{2Z_1 Z_2}{Z_1 + Z_2}$, $P_Z(z)$, is given by

$$P_Z(z) = 1 - z \sqrt{\lambda_1 \lambda_2} e^{-\frac{z}{2}(\lambda_1 + \lambda_2)} K_1(z \sqrt{\lambda_1 \lambda_2}). \tag{36}$$

Thus, the CCDF of Z can be written as

$$1 - P_Z(z) = z \sqrt{\lambda_1 \lambda_2} e^{-\frac{z}{2}(\lambda_1 + \lambda_2)} K_1(z \sqrt{\lambda_1 \lambda_2}). \tag{37}$$

Hence, we have

$$\bar{F}_Y^{(3)}(y) = \frac{y}{\xi'} e^{-y\phi'} K_1 \left(\frac{y}{\xi'} \right), \tag{38}$$

where $\xi' = 2d_N\rho\sqrt{\frac{\beta_{R_N D_N} \cdot \beta_{R_1 D_N}}{\beta_{R_N} \cdot \beta_{R_1}}}$ and $\phi' = \frac{1}{4d_N\rho} \left(\frac{b_{R_1}}{\beta_{R_N D_N}} + \frac{b_{R_N}}{\beta_{R_1 D_N}} \right)$. Note that the analysis result of the method is consistent with the extended form of (32), which verifies the correctness of our analysis idea. Therefore, by substituting (38) into (35), the CCDF of Y can finally be expressed as

$$\bar{F}_Y(y) = \frac{y}{\xi'} e^{-y\phi' - \frac{y}{a_N\rho} \left(\frac{1}{\beta_{SR_N}} + \frac{1}{\beta_{SR_1}} \right)} K_1 \left(\frac{y}{\xi'} \right). \quad (39)$$

Letting $\psi = \phi' + \frac{1}{a_N\rho} \left(\frac{1}{\beta_{SR_N}} + \frac{1}{\beta_{SR_1}} \right)$ and employing [31]—Eq. (3.352.2), the result (the achievable ergodic rate) of the case $i = N$ (\tilde{C}_{s_N}) can be obtained as (40).

$$\begin{aligned} \tilde{C}_{s_N} &= \frac{1}{2\xi' \ln 2} \int_0^\infty \frac{y e^{-\psi y} K_1 \left(\frac{y}{\xi'} \right)}{1+y} dy \\ &\sim -\frac{e^\psi}{2 \ln 2} Ei(-\psi). \end{aligned} \quad (40)$$

Finally, considering all the cases of i where $i \in [1, N]$ together, we can describe the ergodic SR of our proposed system in a closed form as (41).

$$\tilde{C}_{sum} \sim \frac{e^\phi}{2 \ln 2} \left[Ei \left(-\phi \frac{a_i}{\tau} - \phi \right) - Ei(-\phi) \right] - \frac{e^\psi}{2 \ln 2} Ei(-\psi). \quad (41)$$

3.2 Outage probability analysis

According to the QoS rates required by users, each user has a predetermined target data rate. When the link capacity cannot meet the required user data rate, communication interruption will occur. Thus, the analysis of outage probability is one of the important performance metrics in our proposed system. In this section, we will analyze the closed-form solutions of the outage probability with its asymptotic expressions. Assume that for s_i the user's target rate is R_{s_i} and the predefined target rate threshold is $R_{\tau_{s_i}}$, the outage probability can be described as

$$\begin{aligned} P_{out} &= 1 - Pr \left\{ R_{s_1} > 2^{2R\tau_{s_1}} - 1, \dots, R_{s_i} > 2^{2R\tau_{s_i}} - 1, \dots, R_{s_N} > 2^{2R\tau_{s_N}} - 1 \right\} \\ &= 1 - Pr \left\{ (R_{s_1} > 2^{2R\tau_{s_1}} - 1) \cap \dots \cap (R_{s_N} > 2^{2R\tau_{s_N}} - 1) \right\} \\ &= 1 - \prod_{i=1}^N \underbrace{Pr(R_{s_i} > 2^{2R\tau_{s_i}} - 1)}_{u_i}. \end{aligned} \quad (42)$$

Assume that $\omega_i = 2^{2R\tau_{s_i}} - 1$ for the sake of simplicity and convenience of analysis, and refer to the above analysis of ergodic SR; therefore, the exact expressions of u_i where

$i \in [1, N - 1]$ can be described as

$$\begin{aligned} u_i &= Pr \left\{ \min(\gamma_{R_i}^{(s_i)}, \gamma_{R_{i+1}}^{(s_i)}, \gamma_{D_i}^{(s_i)}) > \omega_i \right\} \\ &= Pr \left\{ \gamma_{R_i}^{(s_i)} > \omega_i \right\} Pr \left\{ \gamma_{R_{i+1}}^{(s_i)} > \omega_i \right\} Pr \left\{ \gamma_{D_i}^{(s_i)} > \omega_i \right\} \end{aligned} \quad (43)$$

i.e.,

$$u_i = \begin{cases} e^{-\frac{\omega_i}{4d_i\rho} \left(\frac{b_{R_{i+1}}}{\beta_{R_i D_i}} + \frac{b_{R_i}}{\beta_{R_{i+1} D_i}} \right)} & \text{for } \omega_i < \frac{a_i}{\tau} \\ 0 & \text{for } \omega_i > \frac{a_i}{\tau} \end{cases}. \quad (44)$$

For the case u_N , it can be written as

$$\begin{aligned} u_N &= Pr \left\{ \min(\gamma_{R_N}^{(s_N)}, \gamma_{R_1}^{(s_N)}, \gamma_{D_N}^{(s_N)}) > \omega_N \right\} \\ &= Pr \left\{ \gamma_{R_N}^{(s_N)} > \omega_N \right\} Pr \left\{ \gamma_{R_1}^{(s_N)} > \omega_N \right\} Pr \left\{ \gamma_{D_N}^{(s_N)} > \omega_N \right\}. \end{aligned} \quad (45)$$

similarly,

$$u_N = \begin{cases} e^{-\frac{\omega_N}{4d_N\rho} \left(\frac{b_{R_1}}{\beta_{R_N D_N}} + \frac{b_{R_N}}{\beta_{R_1 D_N}} \right) - \frac{\omega_N}{a_N\rho} \left(\frac{1}{\beta_{SR_N}} + \frac{1}{\beta_{SR_1}} \right)} & \text{for } \omega_N < \frac{a_N}{\tau} \\ 0 & \text{for } \omega_N > \frac{a_N}{\tau} \end{cases}. \quad (46)$$

By substituting (44) and (46) back into (35), with the condition $\omega_i < \frac{a_i}{\tau}$ where $i \in [1, N]$, the outage probability can be obtained in the closed-form expression as

$$\begin{aligned} P_{out} &= 1 - \prod_{i=1}^N \underbrace{Pr(R_{s_i} > 2^{2R\tau_{s_i}} - 1)}_{u_i} \\ &= 1 - e^{-\left[\sum_{i=1}^{N-1} \frac{\omega_i}{4d_i\rho} \left(\frac{b_{R_{i+1}}}{\beta_{R_i D_i}} + \frac{b_{R_i}}{\beta_{R_{i+1} D_i}} \right) \right]} \\ &\quad e^{-\left[\frac{\omega_N}{4d_N\rho} \left(\frac{b_{R_1}}{\beta_{R_N D_N}} + \frac{b_{R_N}}{\beta_{R_1 D_N}} \right) - \frac{\omega_N}{a_N\rho} \left(\frac{1}{\beta_{SR_N}} + \frac{1}{\beta_{SR_1}} \right) \right]}. \end{aligned} \quad (47)$$

4 Numerical results

In this section, we examine the performance of our proposed CMRS-NOMA scheme in terms of ergodic SR and outage probability. Taking the number of signals N as 3 as an example, we study the ergodic rates of three signals (s_1, s_2 , and s_3) and the corresponding sum rate through a Monte Carlo simulation, and we compare the obtained analytical results with the simulated ones. Eight rates of both types match better with the increase in transmit SNR, as shown in Figs. 2 and 3. We also investigate the ergodic SR performance with respect to power allocation factors a_i and b_i for our proposed scheme with different transmit SNRs ρ , as seen in Figs. 4, 5, 6, 7, 8, and 9. In addition, we investigate the outage performance of our proposed scheme, as demonstrated in Fig. 10, and also compare the simulations with the analysis results in three cases.

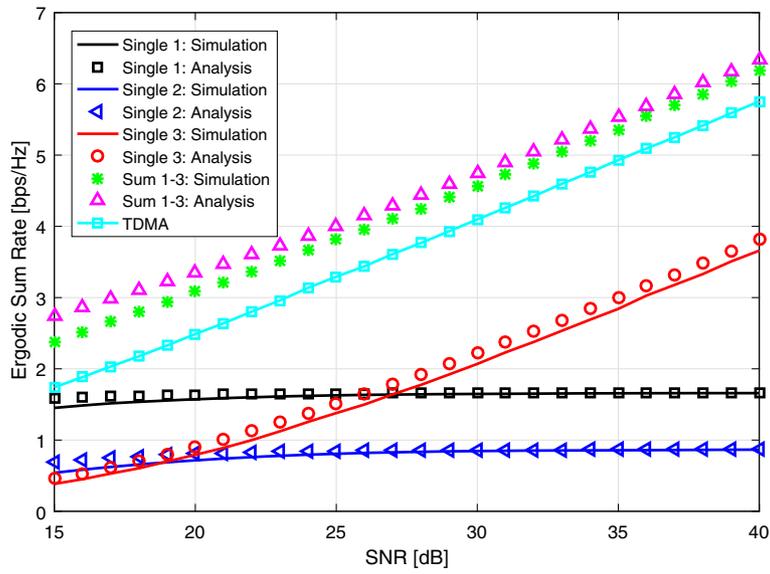


Fig. 2 The ergodic SRs achieved by our proposed MCRS-NOMA scheme with fixed $\nu_{SR1} = 7, \nu_{SR2} = 9, \nu_{SR3} = 10, \nu_{R1D} = 3, \nu_{R2D} = 4,$ and $\nu_{R3D} = 2$

4.1 Ergodic SR

Figure 2 presents the average rates of the proposed algorithm with fixed $\nu_{SR_i}, \nu_{R_iD_j}$ and corresponding $\beta_{SR_i}, \beta_{R_iD_j}$ for the transmission when $a_1 = 0.9, a_2 = 0.07, a_3 = 1 - a_1 - a_2,$ and $b_i = \frac{1}{3} (i \in [1, 3]).$ Clearly, due to the approximation of the analytical formula, at low transmit SNR, the theoretical analysis of the exact average rate is not in perfect agreement with the simulation results.

However, in the high SNR region, the asymptotic results match the simulation results well, which is consistent with the practical situation. In Fig. 3, the power allocation factors are modified, i.e., $a_1 = 0.88, a_2 = 0.09,$ and $a_3 = 1 - a_1 - a_2,$ and the graphic shows the same trend: the simulated results and analytical results are almost identical. Thus, as shown in the figures, the differences of power allocation factors and signal-to-noise ratio have an effect

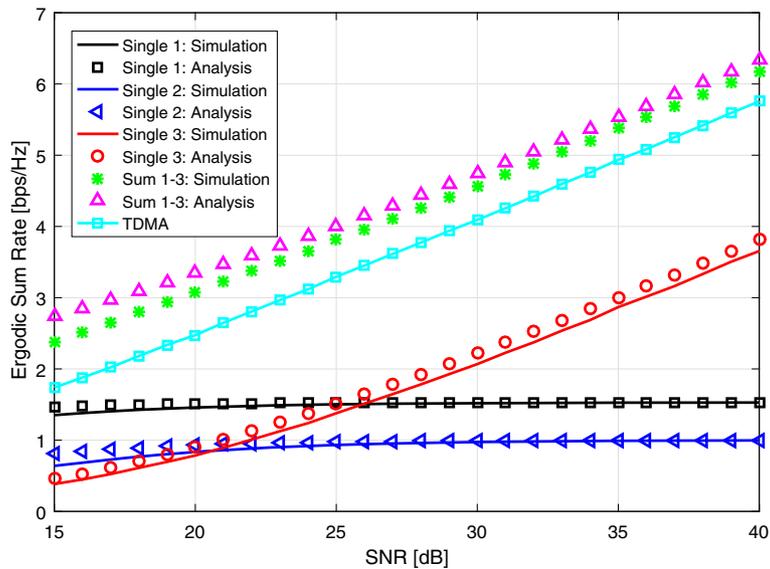


Fig. 3 The ergodic SRs achieved by modifying power allocation factors with fixed $\nu_{SR1} = 7, \nu_{SR2} = 9, \nu_{SR3} = 10, \nu_{R1D} = 3, \nu_{R2D} = 4,$ and $\nu_{R3D} = 2$

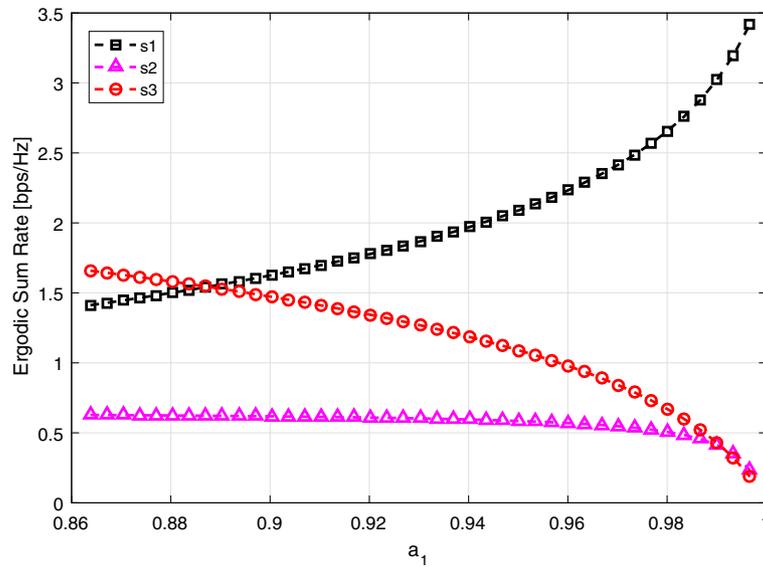


Fig. 4 The ergodic SRs achieved by our proposed scheme versus different power allocation factors a_1 for transmit SNR as $\rho = 25$ dB

on the average rates. In sum, the ergodic SR of our proposed case has an advantage over the TDMA scheme in which the ergodic SR can be described as (48).

$$C_{SR} = \sum_{i=1}^N \frac{1}{2N} \log_2 \left(1 + \frac{1}{2N} \min \{ v_{SR_i} \rho, v_{SR_{i+1}} \rho, v_{R_i D_i} \rho + v_{R_{i+1} D_i} \rho \} \right). \quad (48)$$

Figures 4 and 5 depict the ergodic SR performance with respect to power allocation factor a_1 for our proposed scheme with different transmit SNRs as $\rho = \{25, 30\}$ dB. For

both figures, we have set fixed $v_{SR_i} = 10$ and $v_{R_i D_i} = 2$, and we have a predetermined value range for different power allocation factors, i.e., $a_1 \in [0, 0.997]$, $a_2 = 0.6 * (1 - a_1)$, and $a_3 = 0.4 * (1 - a_1)$. Particularly, to simplify the analysis, we further assume that $b_i = \frac{1}{3} (i \in [1, 3])$, and then $d_i = b_{R_i} * b_{R_{i+1}} * a_i = \frac{1}{9} * a_i (i \in [1, 3])$ will be taken in the simulation. As shown in the figures, an optimal value of a_1 exists that maximizes the ergodic SR. Meanwhile, with the increase in the power ratio a_1 of NOMA symbols,

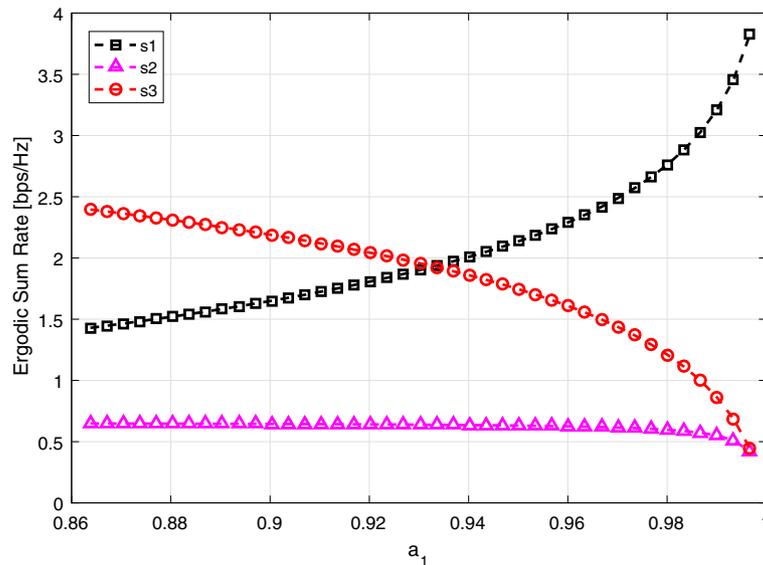


Fig. 5 The ergodic SRs achieved by our proposed scheme versus different power allocation factors a_1 for transmit SNR as $\rho = 30$ dB

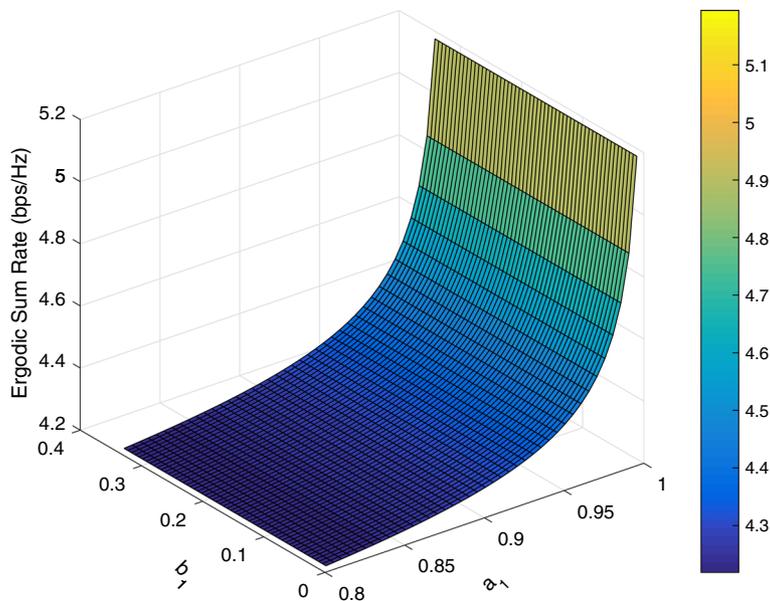


Fig. 6 The ergodic SRs achieved by our proposed scheme versus power allocation factors a_1 and b_1 with fixed $\nu_{SR1} = 3, \nu_{SR2} = 5, \nu_{SR3} = 6, \nu_{R1D} = 10, \nu_{R2D} = 18, \nu_{R3D} = 25$, and $\rho = 25$ dB

the rate of data symbol s_1 increases while those of data symbols s_2 and s_3 decrease. In addition, comparing these figures, the corresponding a_1 for the optimal ergodic SR will be close to 1 with an increasing SNR.

Figures 6 and 7 show the ergodic SR performance of our proposed scheme versus the power allocation factors

a_1 and b_1 with fixed $\nu_{SR1} = 3, \nu_{SR2} = 5, \nu_{SR3} = 6, \nu_{R1D} = 10, \nu_{R2D} = 18, \nu_{R3D} = 25$, and $\rho = \{25, 35\}$ dB. We have a predetermined value range for power allocation factors a_i , i.e., $a_1 \in (0.801, 0.999)$, $a_2 = 0.8 * a_1$, and $a_3 = 1 - a_1 - a_2$. The maximum results of the ergodic SR are obtained with a_1 approaching to 1. In Fig. 6,

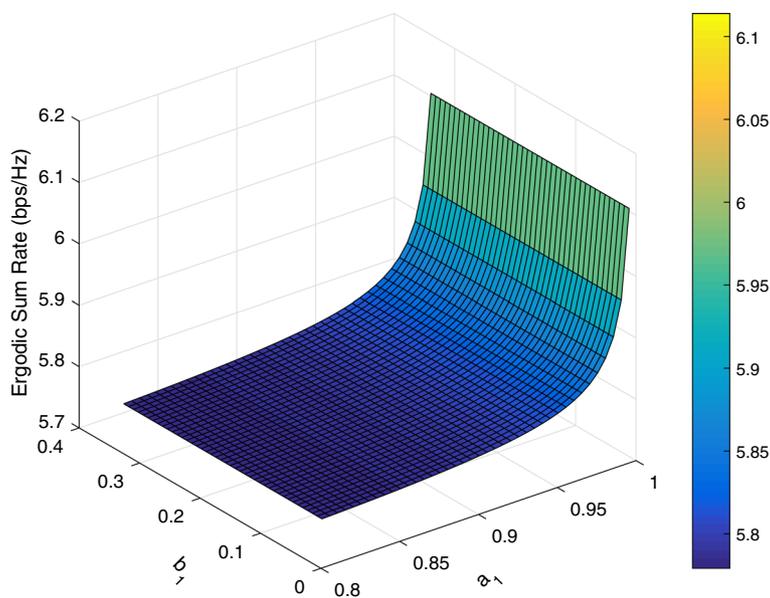


Fig. 7 The ergodic SRs achieved by our proposed scheme versus power allocation factors a_1 and b_1 with fixed $\nu_{SR1} = 3, \nu_{SR2} = 5, \nu_{SR3} = 6, \nu_{R1D} = 10, \nu_{R2D} = 18, \nu_{R3D} = 25$, and $\rho = 35$ dB

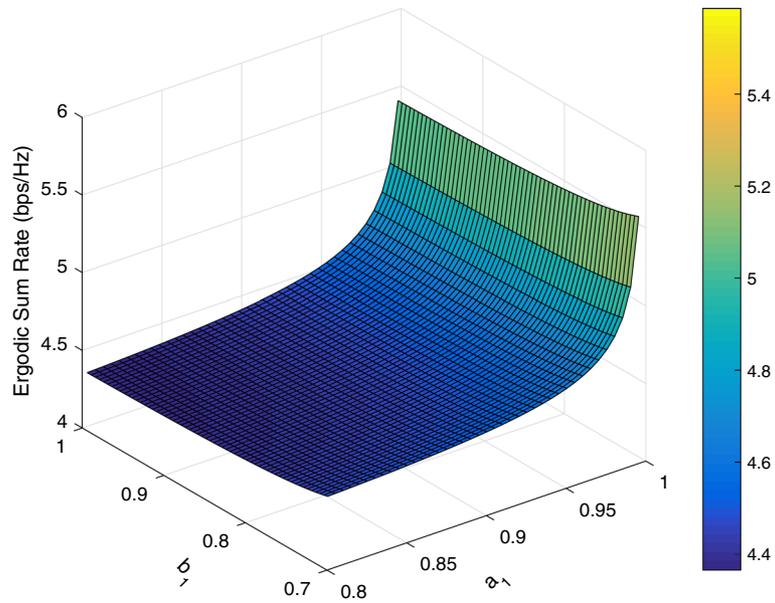


Fig. 8 The ergodic SRs achieved by our proposed scheme versus power allocation factors a_1 and b_1 with fixed $\nu_{SR1} = 3, \nu_{SR2} = 5, \nu_{SR3} = 6, \nu_{R1D} = 10, \nu_{R2D} = 18, \nu_{R3D} = 25$, and $\rho = 25$ dB

the maximum result is 5.194 bps/Hz with a_1 as 0.996, whereas these values are 6.114 bps/Hz and 0.996, respectively, for Fig. 7. As shown, the surfaces of ergodic SR in the above figures are relatively flat. This phenomenon can be explained by assuming that the value of b_i is uniform. Figures 8 and 9 are the results of non-uniform values of b_i , e.g., $b_1 \in (0.701, 0.999), b_2 = 0.7 * b_1$, and $b_3 = 1 - b_1 - b_2$. Hence, the surface of ergodic SR presents a process of

changes in values. Accordingly, the maximum value in Fig. 8 is 5.586 bps/Hz with b_1 as 0.701, which is beyond the range shown in Fig. 6. By using the data cursor in Fig. 9, we can observe that the maximum SR is 6.338 bps/Hz.

4.2 Outage probability

Figure 10 demonstrates the outage performance of our proposed scheme in terms of the simulation and analytical

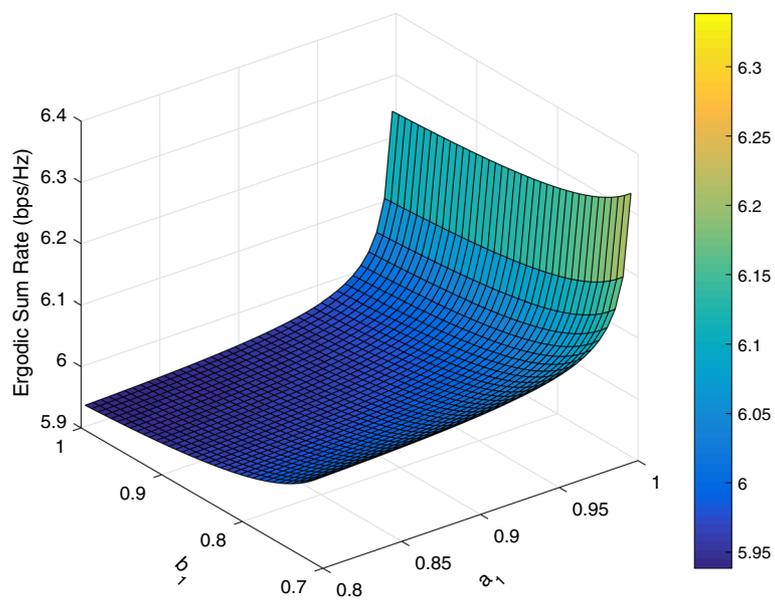


Fig. 9 The ergodic SRs achieved by our proposed scheme versus power allocation factors a_1 and b_1 with fixed $\nu_{SR1} = 3, \nu_{SR2} = 5, \nu_{SR3} = 6, \nu_{R1D} = 10, \nu_{R2D} = 18, \nu_{R3D} = 25$, and $\rho = 35$ dB

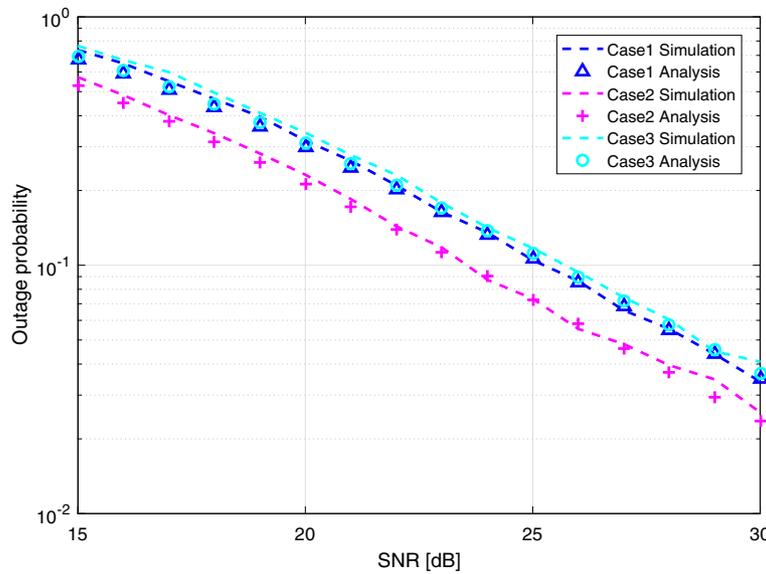


Fig. 10 The outage probability for our proposed scheme with $\rho = \{15, 30\}$ dB in three cases. Case 1 versus the target rate as 0.6 with fixed $\nu_{SR1} = 10$, $\nu_{SR2} = 18$, $\nu_{SR3} = 25$, $\nu_{R1D} = 3$, $\nu_{R2D} = 3$, and $\nu_{R3D} = 3$. Case 2 versus the target rate as 0.6 with fixed $\nu_{SR1} = 10$, $\nu_{SR2} = 18$, $\nu_{SR3} = 25$, $\nu_{R1D} = 5$, $\nu_{R2D} = 5$, and $\nu_{R3D} = 5$. Case 3 versus the target rate as 0.8 with fixed $\nu_{SR1} = 10$, $\nu_{SR2} = 18$, $\nu_{SR3} = 25$, $\nu_{R1D} = 5$, $\nu_{R2D} = 5$, and $\nu_{R3D} = 5$

results with three cases, where the power allocation factors are $a_1 = 0.88$, $a_2 = 0.09$, $a_3 = 1 - a_1 - a_2$, and $b_i = \frac{1}{3}$ ($i \in [1, 3]$). The following system setups are considered: (1) $\omega = 0.6$, $\rho = \{15, 30\}$ dB, $\nu_{SR1} = 10$, $\nu_{SR2} = 18$, $\nu_{SR3} = 25$, and $\nu_{R1D} = \nu_{R2D} = \nu_{R3D} = 3$ for case 1; (2) $\omega = 0.6$, $\rho = \{15, 30\}$ dB, $\nu_{SR1} = 10$, $\nu_{SR2} = 18$, $\nu_{SR3} = 25$, and $\nu_{R1D} = \nu_{R2D} = \nu_{R3D} = 5$ for case 2; and (3) $\omega = 0.8$, $\rho = \{15, 30\}$ dB, $\nu_{SR1} = 10$, $\nu_{SR2} = 18$, $\nu_{SR3} = 25$, and $\nu_{R1D} = \nu_{R2D} = \nu_{R3D} = 5$ for case 3. The results reveal that a good match exists between the simulation and analytical results. The outage probabilities of the above three cases change with the transmission SNR (ρ). For example, the values are respectively $\{0.2997, 0.2118, 0.3110\}$, where $\rho = 20$ dB and $\{0.1065, 0.0724, 0.1111\}$ where $\rho = 25$ dB. It is clear that the performance of the outage probability improves as the channel gains increase and the threshold decreases. The former can be observed by comparing case 1 with case 2, while the latter can be shown by comparing case 2 with case 3. The values of the previous examples reveal this rule well.

5 Conclusions and future work

In this paper, a new transmission model, CMRS-NOMA, in cellular networks has been proposed. In the proposed scheme, SIC is employed to decode the receptions from the source to multiple relay nodes, and the exact and asymptotic expressions for the achievable SR and outage probability of the proposed system over the independent Rayleigh fading channels are presented. In addition, we qualitatively analyze the complexity of the multi-stage

power allocation strategy. It is easy to observe that all the signals can be simply and practically decoded. Numerical results have been presented to corroborate the theoretical analyses, and the results have shown that the performance of the ergodic SR for the proposed CMRS-NOMA case gains a significant improvement and outperforms the TDMA. For a high SNR case, the proposed CMRS-NOMA technology can also provide a greater spectral efficiency. Even that the NOMA systems over Rician fading channels and/or Nakagami- m fading channels are more interesting and challenging, our paper considers providing the mentality to study the receiver strategies for cooperative relaying cellular networks. Therefore, investigating different channel models and efficient power allocation methods for the receiving nodes with different QoS requirements are remained as the future works.

Abbreviations

AWGN: Additive white Gaussian noise; BS: Base station; CCDF: Complementary cumulative distribution function; CDF: Cumulative distribution function; CMRS-NOMA: Cooperative multi-relay system with non-orthogonal multiple access; CRS-NOMA: Cooperative relaying system with NOMA; CSI: Channel state information; CU: Corresponding user; IoT: Internet of Things, MIMO: Multiple-input multiple-output, mmWave: Millimeter-wave, MTC: Machine-type communications; NPL: Node pair list; OMA: Orthogonal multiple access; QoS: Quality of services; SC: Superposition coding; SIC: Successive interference cancellation; SINR: Signal-to-interference-plus-noise ratio; SM: Spatial modulation; SR: Sum rate; SWIPT: Simultaneous wireless information and power transfer; TDMA: Time-division multiple access; V2V: Vehicle-to-vehicle

Acknowledgements

The author would like to thank the Editor and reviewers for their valuable and insightful comments. The authors would gratefully acknowledge the grants from the National Natural Science Foundation of China (61801249, 61501264 and 61371113), Nantong University-Nantong Joint Research Center for

Intelligent Information Technology (KFKT2016B01, KFKT2017B01), Basic Scientific Research of Nantong Science and Technology Project (JCZ18052), Natural Science Research Program of Nantong Vocational University (18ZK01), the open research fund of National Mobile Communications Research Laboratory, Southeast University (No.2015D02).

Availability of data and materials

The authors declare that all the data and materials in this manuscript are available.

Authors' contributions

JJ, WD, and GZ conceived and designed the study. JJ, WD, and QS performed the simulations. JJ and LJ wrote the paper. JJ, GZ, QS, LJ, and WD reviewed and edited the manuscript. All authors read and approved the manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹School of Information Science and Technology, Nantong University, Seyuan Road, Nantong, China. ²School of Electronic and Information Engineering, Nantong Vocational University, Qingnian Road, Nantong, China. ³The Nantong Research Institute for Advanced Communication Technologies, Chongchuan Road, Nantong, China.

Received: 19 October 2018 Accepted: 21 February 2019

Published online: 18 March 2019

References

- Z. Ding, M. Peng, H. V. Poor, Cooperative non-orthogonal multiple access in 5G systems. *IEEE Commun. Lett.* **19**, 1462–1465 (2015)
- J. B. Kim, I. H. Lee, Capacity analysis of cooperative relaying systems using non-orthogonal multiple access. *IEEE Commun. Lett.* **19**, 1949–1952 (2015)
- M. Wen, E. Basar, Q. Li, B. Zheng, M. Zhang, Multiple-mode orthogonal frequency division multiplexing with index modulation. *IEEE Trans. Commun.* **65**, 3892–3906 (2017)
- M. Wen, B. Ye, E. Basar, Q. Li, F. Ji, Enhanced orthogonal frequency division multiplexing with index modulation. *IEEE Trans. Wirel. Commun.* **16**, 4786–4801 (2017)
- M. Wen, X. Cheng, M. Ma, B. Jiao, H. V. Poor, On the achievable rate of OFDM with index modulation. *IEEE Trans. Signal Process.* **64**, 1919–1932 (2016)
- J. Li, M. Wen, X. Cheng, Y. Yan, S. Song, M. H. Lee, Generalized precoding-aided quadrature spatial modulation. *IEEE Trans. Veh. Technol.* **66**, 1881–1886 (2017)
- F. Cheng, H. Yan, J. Song, V. Leung, Energy-efficient resource allocation for downlink non-orthogonal multiple access network. *IEEE Trans. Commun.* **64**, 3722–3732 (2016)
- Y. Sun, D. Ng, Z. Ding, R. Schober, Optimal joint power and subcarrier allocation for full-duplex multicarrier non-orthogonal multiple access systems. *IEEE Trans. Commun.* **65**, 1077–1091 (2017)
- Y. Saito, A. Benjebbour, Y. Kishiyama, T. Nakamura, Generalized precoding-aided quadrature spatial modulation. *Proc. IEEE Annu. Symp. PIMRC.* **66**, 611–615 (2013)
- M. Xu, M. Wen, F. Ji, W. Duan, Novel receiver design for the cooperative relaying system with non-orthogonal multiple access. *IEEE Commun. Lett.* **66**, 1679–1682 (2016)
- Y. Liu, Z. Ding, M. Elkashlan, H. V. Poor, Cooperative non-orthogonal multiple access with simultaneous wireless information and power transfer. *IEEE J. Sel. Areas Commun.* **34**, 938–953 (2016)
- Z. Yang, Z. Ding, P. Fan, N. Al-Dhahir, A general power allocation scheme to guarantee quality of service in downlink and uplink noma systems. *IEEE Trans. Wirel. Commun.* **15**, 7244–7257 (2016)
- R. Jiao, L. Dai, J. Zhang, R. MacKenzie, M. Hao, On the performance of NOMA-based cooperative relaying systems over Rician fading channels. *IEEE Trans. Veh. Technol.* **66**, 11409–11413 (2017)
- X. Yue, Y. Liu, S. Kang, A. Nallanathan, Performance analysis of NOMA with fixed gain relaying over Nakagami-m fading channels. *IEEE Access.* **5**, 5445–5454 (2017)
- H. Xie, F. Gao, S. Jin, J. Fang, Y. C. Liang, Channel estimation for TDD/FDD massive MIMO systems with channel covariance computing. *IEEE Trans. Wirel. Commun.* **17**, 4206–4218 (2018)
- H. Xie, F. Gao, S. Zhang, S. Jin, A unified transmission strategy for TDD/FDD massive MIMO systems with spatial basis expansion model. *IEEE Trans. Veh. Technol.* **66**, 3170–3184 (2017)
- B. Zheng, M. Wen, E. Basar, F. Chen, Multiple-input multiple-output OFDM with index modulation: low-complexity detector design. *IEEE Trans. Signal Process.* **65**, 2758–2772 (2017)
- J. Zhang, L. Dai, S. Sun, Z. Wang, On the spectral efficiency of massive MIMO systems with low-resolution ADCs. *IEEE Commun. Lett.* **20**, 842–845 (2016)
- J. Choi, Minimum power multicast beamforming with superposition coding for multiresolution broadcast and application to NOMA systems. *IEEE Trans. Commun.* **63**, 791–800 (2015)
- M. F. Hanif, Z. Ding, T. Ratnarajah, G. K. Karagiannidis, A minorization-maximization method for optimizing sum rate in the downlink of non-orthogonal multiple access systems. *IEEE Trans. Signal Process.* **64**, 76–88 (2016)
- Q. Sun, S. Han, C. L. I, Z. Pan, On the ergodic capacity of MIMO NOMA systems. *IEEE Wirel. Commun. Lett.* **4**, 405–408 (2015)
- J. Choi, On the power allocation for MIMO-NOMA systems with layered transmissions. *IEEE Trans. Wirel. Commun.* **15**, 3226–3237 (2016)
- Z. Ding, L. Dai, H. V. Poor, MIMO-noma design for small packet transmission in the internet of things. *IEEE Access.* **4**, 1393–1405 (2016)
- W. Feng, Y. Wang, D. Lin, N. Ge, J. Lu, S. Li, When mmWave communications meet network densification: a scalable interference coordination perspective. *IEEE J. Sel. Areas Commun.* **35**, 1459–1471 (2017)
- B. Wang, F. Gao, S. Jin, H. Lin, G. Y. Li, Spatial- and frequency-wideband effects in millimeter-wave massive MIMO systems. *IEEE Trans. Signal Process.* **66**, 3393–3406 (2018)
- Z. Ding, P. Fan, H. V. Poor, Random beamforming in millimeter-wave noma networks. *IEEE Access.* **5**, 7667–7681 (2017)
- R. Zhang, Z. Zhong, J. Zhao, B. Li, K. Wang, Channel measurement and packet-level modeling for V2I spatial multiplexing uplinks using massive MIMO. *IEEE Trans. Veh. Technol.* **65**, 7831–7843 (2016)
- P. Harris, Performance characterization of a real-time massive MIMO system with LOS mobile channels. *IEEE J. Sel. Areas Commun.* **35**, 1244–1253 (2017)
- Y. Chen, L. Wang, Y. Ai, B. Jiao, L. Hanzo, Performance analysis of NOMA-SM in vehicle-to-vehicle massive MIMO channels. *IEEE J. Sel. Areas Commun.* **35**, 2653–2666 (2017)
- I. S. Gradshteyn, I. M. Ryzhik, *Table of Integrals, Series, and Products*. 7th ed. (Academic, San Diego, 2007)
- M. Abramowitz, I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. 9th ed. (Dover Publications, New York, 1970)
- A. Milton, I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. 10th ed. (US GPO, New York, 1972)