

Optimization of Rate Allocation and Power Control for Rate Splitting Multiple Access (RSMA)

Zhaohui Yang, Mingzhe Chen, Walid Saad, *Fellow, IEEE*, and Mohammad Shikh-Bahaei, *Senior Member, IEEE*

Abstract—In this paper, the sum-rate maximization problem is studied for wireless networks that use downlink rate splitting multiple access (RSMA). In the considered model, the base station (BS) divides the messages that can be transmitted to its users into a “private” part and a “common” part. Here, the common message is a message that multiple users want to receive and the private message is a message that is dedicated to only a specific user. The RSMA mechanism enables a BS to adjust the split of common and private messages so as to control the interference by decoding and treating interference as noise and, thus optimizing the data rate of users. To maximize the users’ sum-rate, the network can determine the rate allocation for the common message to meet the rate demand, and adjust the transmit power for the private message to reduce the interference. This problem is formulated as an optimization problem whose goal is to maximize the sum-rate of all users. To solve this nonconvex maximization problem with a single-antenna BS, the optimal power used for transmitting the private message to the users is first obtained in closed form for a given rate allocation and common message power. Based on the optimal private message transmit power, the optimal rate allocation is then derived under a fixed common message transmit power. Subsequently, an iterative algorithm is proposed to obtain a suboptimal solution of common message transmit power. To solve this nonconvex maximization problem with a multiple-antenna BS, a successive convex approximation method is utilized. Simulation results show that the RSMA can achieve up to 15.6% and 21.5% gains in terms of data rate compared to non-orthogonal multiple access (NOMA) and orthogonal frequency-division multiple access (OFDMA), respectively.

Index Terms—Rate splitting multiple access, sum-rate maximization, rate allocation, power control.

I. INTRODUCTION

Driven by the rapid development of advanced multimedia applications such as virtual reality [2], [3], next-generation wireless networks [4] must support high spectral efficiency

This work was supported in part by the U.S. National Science Foundation under Grant CNS-1836802, CNS-1909372, by the EPSRC SENSE Project under Grant EP/P003486/1, by the Fundamental Research Funds for the Central Universities under Grant 2042021kf1030, and by the EPSRC IoSIRE Project under Grant EP/P022723/1. (*Corresponding author: Mingzhe Chen.*)

Part of the material in this paper will be presented in IEEE ICC 2020 [1]. Z. Yang is with the Department of Electronic and Electrical Engineering, University College London, WC1E 6BT London, UK, and also with the Centre for Telecommunications Research, Department of Engineering, King’s College London, WC2R 2LS, UK, Email: zhaohui.yang@ucl.ac.uk.

M. Chen is with the Chinese University of Hong Kong, Shenzhen, China, and also with the Electrical Engineering Department of Princeton University, USA, Email: mzchen00@gmail.com.

W. Saad is with Wireless@VT, Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA, USA, Email: walid-s@vt.edu.

M. Shikh-Bahaei is with the Centre for Telecommunications Research, Department of Engineering, King’s College London, WC2R 2LS, UK, Email: m.sbahaei@kcl.ac.uk.

and massive connectivity. By splitting users in the power domain, non-orthogonal multiple access (NOMA) can simultaneously serve multiple users at the same frequency or time resource [5]–[10]. Consequently, NOMA-based access scheme can achieve higher spectral efficiency than conventional orthogonal multiple access (OMA) [11]–[13]. However, using NOMA, the users must decode all of the interference as they receive the messages [12], which significantly increases the computational complexity needed for signal processing. To solve this problem, the idea of rate splitting multiple access (RSMA) was proposed in [14]–[16]. In RSMA, the message transmitted to the users is divided into a common message and a private message. The common message is a message decoded by multiple users and the private message is a message that only a specific intended user wishes to receive. To receive the common message, the users first decode the interference from other users. In contrast, to receive the private message, the users only consider the interference from other users’ private messages which can be treated as noise. Therefore, adjusting the split of common and private messages can control the computational complexity and the data rate achieved by RSMA. However, implementing RSMA in wireless networks also face several challenges [16] such as the split of common and private message, resource management for effective private message transmission, and synchronization of message transmission.

A. Related Works

Recently, a number of existing works such as in [16]–[26] studied important problems related to RSMA. The work in [16] introduced the challenges and opportunities of using RSMA for multiple input multiple output (MIMO) wireless networks. In [17], the authors proposed a distributed rate splitting method to maximize the data rates of the users. The authors in [18] evaluated the performance of RSMA, NOMA, and space-division multiple access (SDMA) and showed that RSMA achieves better performance than NOMA and SDMA. The authors in [19] investigated the use of linearly-precoded rate-splitting method for simultaneous wireless information and power transfer networks. In [20], the authors used RSMA to maximize the rate of all users in downlink multi-user multiple input single output (MISO) systems under imperfect channel state information at the transmitter.

The energy efficiency problem for RSMA was studied in [21]. The authors in [22] investigated the rate performance of RSMA at millimeter wave. Based on [21] and [22], the work in [23] studied the energy efficiency of the RSMA and NOMA schemes in a millimeter wave downlink transmission

scenario. The use of RSMA is investigated in [24] for a downlink multiuser MISO system with bounded errors in the channel state information at the transmitter. The authors in [25] analyzed the data rate of using RSMA for two-receiver MISO broadcast channel with finite rate feedback. In [26], the spectral and energy efficiency were investigated for RSMA to non-orthogonal unicast and multicast transmissions.

Most of the existing works such as in [16]–[26] find the local optimal power control solutions for RSMA in different wireless systems such as MIMO and MISO. Note that successive interference cancellation (SIC) is also considered in both RSMA and NOMA. Due to SIC, the signal which is decoded first should be allocated more power as shown in [27, Eq. (5)], [28, Eq. (2)], [29, Eq. (4)], [30, Eq. (8)], and [31, Sec. II-B]. In RSMA, the common messages is decoded first and then the private message is decoded with SIC. To provide better overall performance, the additional power constraint needs to be considered to ensure that the common message is a stronger signal than the private message, and this power constraint is called the *SIC constraint* hereinafter. However, none of these prior works [16]–[26] considers a SIC constraint for power allocation in RSMA.

B. Contributions

To our best knowledge, this is the first work that investigates the use of RSMA under an explicit SIC constraint. Our key contributions include:

- We formulate the rate allocation and power control problem as an optimization problem whose goal is to maximize the network sum-rate under both rate and SIC constraints in SISO systems. To solve this problem, we first derive a closed-form expression for the optimal transmit power of the private message. Our fundamental analysis shows that, with the exception of one of the users, all users are allocated with the minimum power to maintain the minimum rate demand. Then, we obtain the optimal rate allocation given the transmit power of the common message.
- Given the equal data rate requirement, we provide a closed-form solution for the optimum common message transmission power. For the general unequal data rate requirement, we have proposed an iterative algorithm to find a suboptimal solution with low complexity.
- We formulate the sum-rate maximization problem for MISO RSMA with both rate and SIC constraints. To solve this nonconvex problem, a successive convex approximation (SCA) based algorithm is proposed to obtain a local optimal solution.
- Simulation results show that the optimized RSMA algorithm can achieve up to 15.6% and 21.5% gains in terms of data rate compared to NOMA and orthogonal frequency-division multiple access (OFDMA).

In our conference version [1], we only studied the sum-rate maximization for RSMA in single-input single-output (SISO) systems.

The rest of this paper is organized as follows. The system model and problem formulation are described in Section II.

The solutions for SISO RSMA and MISO RSMA are presented in Sections III and IV, respectively. Simulation results are analyzed in Section V. Conclusions are drawn in Section VI.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

Consider a downlink single-cell network that consists of one BS servicing a set \mathcal{K} of K users using RSMA. In RSMA, the message W_k intended for user k is split into a common part $W_{c,k}$ and a private part $W_{p,k}$ [32]. The common parts $W_{c,1}, \dots, W_{c,K}$ of all users are combined into the common message W_c , which is encoded into the common stream s_0 using a codebook shared by all users. Hence, s_0 is a common stream required to be decoded by all users. The private part $W_{p,k}$, containing the remaining part of the message W_k , is encoded into the private stream s_k for use k .

The transmitted signal x of the BS is expressed as:

$$x = \sqrt{p_0}s_0 + \sum_{k=1}^K \sqrt{p_k}s_k, \quad (1)$$

where p_0 is the transmit power of the common stream s_0 (i.e., the common message W_c) and p_k is the transmit power of the private stream s_k (i.e., the private message $W_{p,k}$) transmitted to user k .

The total received signal at user k can be given by:

$$y_k = \sqrt{h_k}x + n_k = \sqrt{h_k p_0}s_0 + \sum_{j=1}^K \sqrt{h_k p_j}s_j + n_k, \quad (2)$$

where h_k represents the channel gain between user k and the BS and n_k is the additive white Gaussian noise with variance σ^2 . The achievable rate of user k decoding common stream s_0 can be expressed as:

$$c_k = B \log_2 \left(1 + \frac{h_k p_0}{h_k \sum_{j=1}^K p_j + \sigma^2} \right), \quad (3)$$

where B is the bandwidth of the BS. Without loss of generality, the channel gains are sorted in ascending order, i.e., $h_1 \leq h_2 \leq \dots \leq h_K$. To ensure that all users can successfully decode common stream s_0 , the rate of common stream should be chosen as [18]:

$$\begin{aligned} \min_{k \in \mathcal{K}} c_k &= \min_{k \in \mathcal{K}} B \log_2 \left(1 + \frac{p_0}{\sum_{j=1}^K p_j + \frac{\sigma^2}{h_k}} \right) \\ &= B \log_2 \left(1 + \frac{p_0}{\sum_{j=1}^K p_j + \frac{\sigma^2}{\min_{k \in \mathcal{K}} h_k}} \right) \\ &\stackrel{(a)}{=} B \log_2 \left(1 + \frac{p_0}{\sum_{j=1}^K p_j + \frac{\sigma^2}{h_1}} \right) = c_1, \end{aligned} \quad (4)$$

where equality (a) follows from the fact that $h_1 \leq h_2 \leq \dots \leq h_K$.

To successfully implement SIC operation at the receiver, the transmit power of each user must satisfy the following constraint [27]:

$$h_k p_0 - h_k \sum_{j=1}^K p_j - \sigma^2 \geq \theta, \quad \forall k \in \mathcal{K}, \quad (5)$$

where θ is the minimum difference between the decoding signal power and the undecoded inter-user interference signal power plus noise power [28]. In (5), $h_k p_0$ means the received common message power at user k , while the term $h_k \sum_{j=1}^K p_j + \sigma^2$ is the power of non-decoded private message of all users (plus noise). According to [33], in order to satisfy SIC in an RSMA-based system, each user should have that the received signal to interference plus noise ratio (SINR) of the common stream signal is larger than its received SINR of the private stream, i.e., the common message is a stronger signal than the private message, as shown in the SIC constraint (5). The SIC constraint (5) is more suitable than the SIC constraint in [33] since our SIC constraint is a linear expression with respect to the transmission power and our SIC constraint can also capture the fact that the received SINR of the common stream signal is larger than the received SINR of the private stream. This minimum difference is required to distinguish the common message to be decoded and the remaining non-decoded private message of all users (plus noise). Based on the channel condition $h_1 \leq h_2 \leq \dots \leq h_K$, constraint (5) can be simplified as:

$$p_0 - \sum_{j=1}^K p_j \geq \frac{\theta + \sigma^2}{h_1}. \quad (6)$$

Given the common stream rate c_1 and the rate a_k allocated to user k , the constraint of each user k 's data rate of receiving common stream is given by:

$$\sum_{k=1}^K a_k \leq c_1, \quad (7)$$

which indicates that the total data rates of all users receiving common stream must be less than the rate c_1 of common stream.

According to (4) and (7), the rate constraint can be further expressed as

$$h_1 p_0 - \left(2^{\frac{\sum_{k=1}^K a_k}{B}} - 1 \right) \left(h_1 \sum_{j=1}^K p_j + \sigma^2 \right) \geq 0. \quad (8)$$

Although both rate constraint (8) and SIC constraint (5) are linear with respect to the transmission power, constraint (8) cannot be equivalently transformed to constraint (5).

After having decoded the common stream s_0 , each user can decode its private stream, the achievable rate of user k decoding its private stream s_k is given by:

$$r_k = B \log_2 \left(1 + \frac{h_k p_k}{h_k \sum_{j=1, j \neq k}^K p_j + \sigma^2} \right). \quad (9)$$

Given the common stream rate a_k and achievable private stream rate r_k , the total transmission rate of user k in RSMA is:

$$r_k^{\text{tot}} = a_k + r_k = a_k + B \log_2 \left(1 + \frac{h_k p_k}{h_k \sum_{j=1, j \neq k}^K p_j + \sigma^2} \right). \quad (10)$$

B. Problem Formulation

Given the considered system model, our objective is to optimize the rate allocation and power control so as to maximize the sum-rate under a total power constraint and individual minimum rate requirements. Mathematically, the sum-rate maximization problem for RSMA can be formulated as:

$$\max_{\mathbf{a}, \mathbf{p}} \sum_{k=1}^K \left(a_k + B \log_2 \left(1 + \frac{h_k p_k}{h_k \sum_{j=1, j \neq k}^K p_j + \sigma^2} \right) \right), \quad (11)$$

$$\text{s.t.} \quad \sum_{k=1}^K a_k \leq B \log_2 \left(1 + \frac{h_1 p_0}{h_1 \sum_{j=1}^K p_j + \sigma^2} \right), \quad (11a)$$

$$a_k + B \log_2 \left(1 + \frac{h_k p_k}{h_k \sum_{j=1, j \neq k}^K p_j + \sigma^2} \right) \geq R_k, \quad \forall k \in \mathcal{K} \quad (11b)$$

$$p_0 - \sum_{j=1}^K p_j \geq \frac{\theta + \sigma^2}{h_1}, \quad (11c)$$

$$\sum_{k=0}^K p_k \leq P, \quad (11d)$$

$$a_k, p_0, p_k \geq 0, \quad \forall k \in \mathcal{K}, \quad (11e)$$

where $\mathbf{a} = [a_1, a_2, \dots, a_K]^T$, $\mathbf{p} = [p_0, p_1, p_2, \dots, p_K]^T$, R_k is the minimum rate demand of user k , and P is the maximum transmit power of the BS. Constraint (11a) ensures that each user can decode the common message. The minimum rate constraints for all users are given in (11b). Constraint (11c) shows the successful SIC power requirement, and (11d) presents the maximum power constraint.

When the rate demand $\sum_{k=1}^K a_k$ is small, we cannot obtain that the common message is a stronger signal than the private message from (11a) [18], [20], [34], [35]. Thus, the SIC constraint (11c) needs to be considered to ensure better overall performance. There is no contradiction between this setting and the flexibility of the RS system. We can first solve two problems: the resource allocation problem with SIC constraint (11c) and the resource allocation problem without a common message. Then, we compare the objective values of these two problems and select the higher objective value.

Since the objective function is not concave, the sum-rate maximization problem (11) is nonconvex. Moreover, the rate and power vectors are coupled in the objective function and constraints, and hence, it is generally hard to solve problem (11).

III. JOINT RATE ALLOCATION AND POWER CONTROL

In this section, we first provide the optimal conditions of problem (11). Then, based on these optimal conditions, the optimal private message transmission power is obtained in closed form under a given rate allocation and common message transmission power. Substituting the optimal private message transmission power in problem (11), the optimal closed-form rate allocation is then derived under a fixed

common message transmission power. Finally, an iterative algorithm is proposed to solve problem (11)

A. Optimal Conditions

Before solving problem (11), we provide some optimal conditions, which will be used to simplify problem (11).

Lemma 1: At the optimal solution $(\mathbf{a}^*, \mathbf{p}^*)$ of problem (11), the common message constraint (11b) holds with equality, i.e., $\sum_{k=1}^K a_k^* = \log_2 \left(1 + \frac{h_1 p_0^*}{h_1 \sum_{j=1}^K p_j^* + \sigma^2} \right)$.

Lemma 1 can be easily proved by contradiction.

Lemma 2: At the optimal solution $(\mathbf{a}^*, \mathbf{p}^*)$ of problem (11), the maximum power constraint (11d) holds with equality, i.e., $\sum_{k=0}^K p_k^* = P$.

Proof: See Appendix A. \square

Applying Lemma 1 and substituting $\sum_{j=1, j \neq k}^K p_j = P - p_0 - p_k$ from Lemma 2 to (11), we can then observe that problem (11) is equivalent to the following problem:

$$\max_{\mathbf{a}, \mathbf{p}} \sum_{k=1}^K a_k + \sum_{k=1}^K B \log_2 \left(\frac{h_k(P - p_0) + \sigma^2}{h_k(P - p_0 - p_k) + \sigma^2} \right), \quad (12)$$

$$\text{s.t.} \quad \sum_{k=1}^K a_k = B \log_2 \left(\frac{h_1 P + \sigma^2}{h_1(P - p_0) + \sigma^2} \right), \quad (12a)$$

$$a_k + B \log_2 \left(\frac{h_k(P - p_0) + \sigma^2}{h_k(P - p_0 - p_k) + \sigma^2} \right) \geq R_k, \forall k \in \mathcal{K}, \quad (12b)$$

$$\sum_{k=0}^K p_k = P, \quad (12c)$$

$$p_0 \geq \frac{P}{2} + \frac{\theta + \sigma^2}{2h_1}, \quad (12d)$$

$$a_k, p_k \geq 0, \quad \forall k \in \mathcal{K}. \quad (12e)$$

To solve problem (12), we can show that it is further equivalent to another optimization problem, which admits a closed-form solution for the optimal private message transmission power.

Lemma 3: The optimal solution of problem (12) is equivalent to the following problem:

$$\max_{\mathbf{a}, \mathbf{p}} B \log_2 \left(\frac{h_1 P + \sigma^2}{h_1(P - p_0) + \sigma^2} \right) + \sum_{k=1}^K B \log_2 \left(\frac{h_k(P - p_0) + \sigma^2}{h_k(P - p_0 - p_k) + \sigma^2} \right), \quad (13)$$

$$\text{s.t.} \quad \sum_{k=1}^K a_k \leq B \log_2 \left(\frac{h_1 P + \sigma^2}{h_1(P - p_0) + \sigma^2} \right), \quad (13a)$$

$$a_k + B \log_2 \left(\frac{h_k(P - p_0) + \sigma^2}{h_k(P - p_0 - p_k) + \sigma^2} \right) \geq R_k, \forall k \in \mathcal{K}, \quad (13b)$$

$$\sum_{k=0}^K p_k = P, \quad (13c)$$

$$p_0 \geq \frac{P}{2} + \frac{\theta + \sigma^2}{2h_1}, \quad (13d)$$

$$p_k \geq 0, 0 \leq a_k \leq R_k, \quad \forall k \in \mathcal{K}. \quad (13e)$$

Proof: See Appendix B. \square

Note that the maximum rate limitation $0 \leq a_k \leq R_k$ is added in constraint (13e), which will prove to be helpful in obtaining the optimal private message transmission power in closed form.

B. Optimal Private Message Transmission Power

Given the simplified problem in (13), next, we find the optimal private message transmission power. Given rate allocation \mathbf{a} and common message power control p_0 , problem (13) becomes

$$\max_{\mathbf{p}} \sum_{k=1}^K B \log_2 \left(\frac{h_k(P - p_0) + \sigma^2}{h_k(P - p_0 - p_k) + \sigma^2} \right), \quad (14)$$

$$\text{s.t.} \quad \sum_{k=1}^K p_k = P - p_0, \quad (14a)$$

$$p_k \geq p_k^{\min}, \quad \forall k \in \mathcal{K}, \quad (14b)$$

where $\bar{\mathbf{p}} = [p_1, p_2, \dots, p_K]$ is a vector of power that is allocated to each user for receiving private message and

$$p_k^{\min} = \left(1 - 2^{-\frac{a_k - R_k}{B}} \right) \left(P - p_0 + \frac{\sigma^2}{h_k} \right). \quad (15)$$

Due to constraint (13e), p_k^{\min} is always non-negative.

Note that the first term in objective function (13) is a constant with given common message power control p_0 , thus the first term in objective function (13) is omitted in problem (14). In (14b), p_k^{\min} is used to meet the minimum rate constraint in (13b), and problem (14) is feasible if and only if $\sum_{k=1}^K p_k^{\min} \leq P - p_0$, which can be given as:

$$\sum_{k=1}^K \left(1 - 2^{-\frac{a_k - R_k}{B}} \right) \left(P - p_0 + \frac{\sigma^2}{h_k} \right) \leq P - p_0. \quad (16)$$

Since objective function is convex, we can infer that the maximization problem (14) is nonconvex. To effectively solve problem (14), the following theorem is presented.

Theorem 1: For the optimal solution $\bar{\mathbf{p}}^*$ of problem (14), there exists one k such that $p_k^* = P - p_0 - \sum_{j=1, j \neq k}^K p_j^{\min}$ and $p_j^* = p_j^{\min}, \forall j \in \mathcal{K}, j \neq k$.

Proof: See Appendix C. \blacksquare

From Theorem 1, the structure of the optimal solution of problem (14) is revealed. Although problem (14) is nonconvex, the optimal solution can be obtained in closed form, which can be given by the following theorem.

Theorem 2: For nonconvex problem (14), the optimal power allocation $\bar{\mathbf{p}}^*$ is

$$p_k^* = P - p_0 - \sum_{j=1, j \neq k}^K \left(1 - 2^{-\frac{a_j - R_j}{B}} \right) \left(P - p_0 + \frac{\sigma^2}{h_j} \right), \quad (17)$$

$$p_j^* = \left(1 - 2^{-\frac{a_j - R_j}{B}} \right) \left(P - p_0 + \frac{\sigma^2}{h_j} \right), \quad \forall j \in \mathcal{K}, j \neq k, \quad (18)$$

and the optimal sum-rate of private message is

$$B \log_2 \left(\frac{P - p_0 + \frac{\sigma^2}{h_k}}{\sum_{j=1, j \neq k}^K \left(1 - 2^{\frac{a_j - R_j}{B}}\right) \left(P - p_0 + \frac{\sigma^2}{h_j}\right) + \frac{\sigma^2}{h_k}} \right) + \sum_{j=1, j \neq k}^K (R_j - a_j), \quad (19)$$

where

$$k = \arg \min_{j \in \mathcal{K}} 2^{\frac{a_j - R_j}{B}} \left(P - p_0 + \frac{\sigma^2}{h_j}\right). \quad (20)$$

Proof: See Appendix D. \blacksquare

Theorem 2 states that it is optimal for the BS to allocate more power to the user that can maximize the sum-rate while allocating the minimum transmit power that can meet the data rate requirement for all other users.

For the special case with $a_j = R_j, \forall j \in \mathcal{K}$, we can obtain $p_j^{\min} = 0$ and $k = \arg \min_{j \in \mathcal{K}} \frac{\sigma^2}{h_j} = K$ according to (20), i.e., all the power should be allocated to the user with the highest channel gain. This observation is trivial since allocating the maximum power to the user with the highest channel gain will always improve the rate.

For the special case in which $\mathbf{a} = \mathbf{0}$, i.e., the broadcast channel without SIC, we have the following corollary that follows from Theorem 2.

Corollary 1: For the downlink nonconvex sum-rate maximization problem in broadcast channel, which is given by:

$$\max_{\mathbf{p}} \sum_{k=1}^K B \log_2 \left(1 + \frac{h_k p_k}{h_k \sum_{j=1, j \neq k}^K p_j + \sigma^2} \right), \quad (21)$$

$$\text{s.t. } B \log_2 \left(1 + \frac{h_k p_k}{h_k \sum_{j=1, j \neq k}^K p_j + \sigma^2} \right) \geq R_k, \forall k \in \mathcal{K}, \quad (21a)$$

$$\sum_{k=1}^K p_k \leq P, \quad (21b)$$

$$p_k \geq 0, \quad \forall k \in \mathcal{K}, \quad (21c)$$

the maximization problem in (21) is feasible if and only if

$$\sum_{j=1}^K \left(1 - 2^{\frac{-R_j}{B}}\right) \left(P - p_0 + \frac{\sigma^2}{h_j}\right) \leq P - p_0, \quad (22)$$

and the optimal power allocation \mathbf{p}^* is to meet the minimum rate requirements of all users except one user, i.e.,

$$p_k^* = P - \sum_{j=1, j \neq k}^K \left(1 - 2^{\frac{-R_j}{B}}\right) \left(P + \frac{\sigma^2}{h_j}\right), \quad (23)$$

$$p_j^* = \left(1 - 2^{\frac{-R_j}{B}}\right) \left(P + \frac{\sigma^2}{h_j}\right), \quad \forall j \in \mathcal{K}, j \neq k, \quad (24)$$

and the optimal objective value is

$$B \log_2 \left(\frac{P + \frac{\sigma^2}{h_k}}{\sum_{j=1, j \neq k}^K \left(1 - 2^{\frac{-R_j}{B}}\right) \left(P + \frac{\sigma^2}{h_j}\right) + \frac{\sigma^2}{h_k}} \right) + \sum_{j=1, j \neq k}^K R_j, \quad (25)$$

where

$$k = \arg \min_{j \in \mathcal{K}} 2^{\frac{-R_j}{B}} \left(P + \frac{\sigma^2}{h_j}\right). \quad (26)$$

Corollary 1 provides the optimal power allocation that maximizes the downlink sum-rate of the broadcast channel. From this result, we can see that more power should be allocated to one of the users compared to allocating the minimum transmit power to all other users. The user that should be allocated more power is jointly determined by the channel gain and the minimum rate demand, as shown in (26).

C. Optimal Rate Allocation

In the previous subsection, the optimal power allocation vector $\bar{\mathbf{p}}$ can be obtained as a function of the rate allocation vector \mathbf{a} and common message power p_0 . Thus, substituting the optimal power allocation vector $\bar{\mathbf{p}}$ given in (17) and (18) in Theorem 2, the original problem in (13) can be simplified as:

$$\max_{\mathbf{a}, p_0} B \log_2 \left(\frac{h_1 P + \sigma^2}{h_1 (P - p_0) + \sigma^2} \right) + \sum_{j=1, j \neq k}^K (R_j - a_j) + B \log_2 \left(\frac{P - p_0 + \frac{\sigma^2}{h_k}}{\sum_{j=1, j \neq k}^K \left(1 - 2^{\frac{a_j - R_j}{B}}\right) \left(P - p_0 + \frac{\sigma^2}{h_j}\right) + \frac{\sigma^2}{h_k}} \right), \quad (27)$$

$$\text{s.t. } \sum_{j=1}^K a_j \leq B \log_2 \left(\frac{h_1 P + \sigma^2}{h_1 (P - p_0) + \sigma^2} \right), \quad (27a)$$

$$k = \arg \min_{j \in \mathcal{K}} 2^{\frac{a_j - R_j}{B}} \left(P - p_0 + \frac{\sigma^2}{h_j}\right), \quad (27b)$$

$$\sum_{j=1}^K \left(1 - 2^{\frac{a_j - R_j}{B}}\right) \left(P - p_0 + \frac{\sigma^2}{h_j}\right) \leq P - p_0, \quad (27c)$$

$$p_0 \geq \frac{P}{2} + \frac{\theta + \sigma^2}{2h_1}, \quad (27d)$$

$$0 \leq a_j \leq R_j, \quad \forall j \in \mathcal{K}, \quad (27e)$$

where constraint (27b) is added since more power should be allocated to user k to maximize the sum-rate while other users are allocated the minimum power to ensure the minimum rate constraint. Constraint (27c) follows from (16), which ensures that the private message power control problem is feasible.

Due to objective function and constraints (27a)-(27c), problem (27) is nonconvex and, hence, it is generally hard to directly optimize rate allocation \mathbf{a} and private power p_0 . To solve problem (27), we first fix the private message transmission power and derive the optimal rate allocation. Given common message transmission power p_0 , problem (27) becomes

$$\max_{\mathbf{a}} \sum_{j=1, j \neq k}^K (R_j - a_j) + B \log_2 \left(\frac{P - p_0 + \frac{\sigma^2}{h_k}}{\sum_{j=1, j \neq k}^K \left(1 - 2^{\frac{a_j - R_j}{-B}}\right) \left(P - p_0 + \frac{\sigma^2}{h_j}\right) + \frac{\sigma^2}{h_k}} \right), \quad (28)$$

$$\text{s.t.} \quad \sum_{j=1}^K a_j \leq c_1, \quad (28a)$$

$$k = \arg \min_{j \in \mathcal{K}} 2^{\frac{a_j - R_j}{-B}} \left(P - p_0 + \frac{\sigma^2}{h_j}\right), \quad (28b)$$

$$\sum_{j=1}^K \left(1 - 2^{\frac{a_j - R_j}{-B}}\right) \left(P - p_0 + \frac{\sigma^2}{h_j}\right) \leq P - p_0, \quad (28c)$$

$$0 \leq a_j \leq R_j, \quad \forall j \in \mathcal{K}. \quad (28d)$$

Despite the nonconvexity of (28), the optimal solution of problem (28) can be precisely formulated in the following theorem. Before presenting the theorem, rearrange users in the descending order $\pi_1, \pi_2, \dots, \pi_K$ of $2^{\frac{-R_j}{-B}} \left(P - p_0 + \frac{\sigma^2}{h_j}\right)$, i.e.,

$$2^{\frac{-R_{\pi_l}}{-B}} \left(P - p_0 + \frac{\sigma^2}{h_{\pi_l}}\right) \geq 2^{\frac{-R_{\pi_{l+1}}}{-B}} \left(P - p_0 + \frac{\sigma^2}{h_{\pi_{l+1}}}\right), \quad \forall j < l. \quad (29)$$

Theorem 3: The optimal solution \mathbf{a}^* of problem (28) is:

i) if $c_1 \leq \sum_{j=1}^K R_j$,

$$a_{\pi_j}^* = \begin{cases} R_{\pi_j}, & \text{if } j < l, \\ c_1 - \sum_{m=1}^{j-1} R_{\pi_m}, & \text{if } j = l, \\ 0, & \text{otherwise,} \end{cases} \quad (30)$$

where l satisfies $\sum_{j=1}^{l-1} R_j \leq c_1 < \sum_{j=1}^l R_j$.

ii) if $c_1 \geq \sum_{j=1}^K R_j$,

$$a_{\pi_j}^* = R_j, \quad \forall j \in \mathcal{K}. \quad (31)$$

Proof: See Appendix E. ■

Theorem 3 provides the optimal rate allocation of problem (28) in closed form. According to Theorem 3, we can determine the user that the BS will allocate additional power using the following lemma.

Lemma 4: It is optimal to allocate the additional power to user π_K , i.e.,

$$\begin{aligned} k &= \arg \min_{j \in \mathcal{K}} 2^{\frac{a_j - R_j}{-B}} \left(P - p_0 + \frac{\sigma^2}{h_j}\right) \\ &= \arg \min_{j \in \mathcal{K}} 2^{\frac{-R_{\pi_j}}{-B}} \left(P - p_0 + \frac{\sigma^2}{h_{\pi_j}}\right) = \pi_K. \end{aligned} \quad (32)$$

D. Power Control Optimization

With the optimal rate allocation given in Theorem 3, it remains to optimize common message transmission power p_0 of problem (27).

Theorem 4: Given the equal data rate requirement, i.e., $R_1 = \dots = R_K = R$, the optimal common message transmission power p_0^* of problem (27) satisfies:

$$p_0^* \in \left\{ \frac{P}{2} + \frac{\theta + \sigma^2}{2h_1}, (1 - 2^{-R}) \left(P + \frac{\sigma^2}{h_1}\right), \dots, (1 - 2^{-KR}) \left(P + \frac{\sigma^2}{h_1}\right), P_1, \dots, P_K \right\}, \quad (33)$$

where P_k satisfies

$$\begin{aligned} & \left(h_1(P - P_k) + \sigma^2 - (h_1 P + \sigma^2) 2^{\frac{-R}{-B}}\right) \left(P - P_k + \frac{\sigma^2}{h_k}\right) \\ & + \left(1 - 2^{\frac{-R}{-B}}\right) \sum_{j=k}^K (h_1(P - P_k) + \sigma^2) \left(P - p_0^* + \frac{\sigma^2}{h_j}\right) \\ & = (P - p_0^*)(h_1(P - P_k) + \sigma^2). \end{aligned} \quad (34)$$

Proof: See Appendix F. ■

Theorem 4 indicates that the optimal common message transmission power p_0^* lies in a finite space of potential candidates. The optimal solution p_0^* of problem (27) will be one of the solutions with the highest objective value. Since equation (34) is a quadratic equation, which can be effectively solved via the quadratic root formula.

For the general unequal data rate requirement, it is hard to obtain the optimal rate allocation and power control of problem (27). To obtain a suboptimal solution of problem (27), we propose an iterative algorithm through obtaining the optimal rate allocation with given transmission power and updating the optimal transmission power with optimized rate allocation, as shown in Algorithm 1.

Theorem 5: Given rate allocation \mathbf{a} , the optimal common message transmission power p_0^* of problem (27) satisfies:

$$p_0^* \in \left\{ \frac{P}{2} + \frac{\theta + \sigma^2}{2h_1}, \left(1 - 2^{-\sum_{k=1}^K a_k}\right) \left(P + \frac{\sigma^2}{h_1}\right), P_0 \right\}, \quad (35)$$

where P_0 is the unique solution to

$$\sum_{j=1}^K \left(1 - 2^{\frac{a_j - R_j}{-B}}\right) \left(P - P_0 + \frac{\sigma^2}{h_j}\right) = P - P_0. \quad (36)$$

Proof: See Appendix G. ■

In each step of Algorithm 1, the main complexity lies in solving (28) given p_0 . The complexity of solving the problem in (28) is $\mathcal{O}(K)$ according to Theorem 3 and the complexity of Algorithm 1 is $\mathcal{O}(LK)$, where L denotes the number of iterations.

Algorithm 1 Iterative Rate Allocation and Power Control

- 1: Initialize $\mathbf{a}^{(0)}$, $p_0^{(0)}$, and iteration number $t = 1$.
 - 2: **repeat**
 - 3: Given $p_0^{(t-1)}$, obtain the optimal $\mathbf{a}^{(t)}$ of problem (27) according to Theorem 3.
 - 4: Given $\mathbf{a}^{(t)}$, obtain the optimal $p_0^{(t)}$ of problem (27) according to Theorem 5.
 - 5: Set $t = t + 1$.
 - 6: **until** the objective value (27) converges.
-

IV. DOWNLINK SUM-RATE MAXIMIZATION FOR MISO RSMA

A. Problem Formulation

We consider the case in which the BS is equipped with N antennas. Let the common message of all users be s_0 and the private message of each user k be s_k . The transmitted signal \mathbf{x} of the BS is expressed as:

$$\mathbf{x} = \mathbf{w}_0 s_0 + \sum_{k=1}^K \mathbf{w}_k s_k, \quad (37)$$

where \mathbf{w}_0 is the beamforming vector of the common message s_0 and \mathbf{w}_k is the beamforming vector of the private message s_k transmitted to user k .

The total received message at user k can be given by:

$$y_k = \mathbf{h}_k^H \mathbf{x} + n_k = \mathbf{h}_k^H \mathbf{w}_0 s_0 + \sum_{j=1}^K \mathbf{h}_k^H \mathbf{w}_j s_j + n_k, \quad (38)$$

where \mathbf{h}_k represents the channel coefficients between user k and the BS. The achievable rate of user k decoding common message s_0 can be expressed as:

$$c_k = B \log_2 \left(1 + \frac{|\mathbf{h}_k^H \mathbf{w}_0|^2}{\sum_{j=1}^K |\mathbf{h}_k^H \mathbf{w}_j|^2 + \sigma^2} \right). \quad (39)$$

To ensure that all users can successfully decode common message s_0 , the rate of common message should be chosen as [18] $\min_{k \in \mathcal{K}} c_k$.

To successfully implement the SIC operation at the receiver, the transmit power of each user must satisfy the following constraint [27]:

$$|\mathbf{h}_k^H \mathbf{w}_0|^2 - \sum_{j=1}^K |\mathbf{h}_k^H \mathbf{w}_j|^2 - \sigma^2 \geq \theta, \quad \forall k \in \mathcal{K}. \quad (40)$$

Given the common message rate $\min_{k \in \mathcal{K}} c_k$ and the rate a_k allocated to user k , the constraint of each user k 's data rate of receiving common message is given by:

$$\sum_{k=1}^K a_k \leq c_k, \quad \forall k \in \mathcal{K}. \quad (41)$$

After having decoded the common message s_0 , each user can decode its private message, the achievable rate of user k decoding its private message s_k is given by:

$$r_k = B \log_2 \left(1 + \frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum_{j=1, j \neq k}^K |\mathbf{h}_k^H \mathbf{w}_j|^2 + \sigma^2} \right). \quad (42)$$

Mathematically, the sum-rate maximization problem for

RSMA can be formulated as:

$$\max_{\mathbf{a}, \mathbf{w}} \sum_{k=1}^K \left(a_k + B \log_2 \left(1 + \frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum_{j=1, j \neq k}^K |\mathbf{h}_k^H \mathbf{w}_j|^2 + \sigma^2} \right) \right), \quad (43)$$

$$\text{s.t.} \quad \sum_{k=1}^K a_k \leq B \log_2 \left(1 + \frac{|\mathbf{h}_k^H \mathbf{w}_0|^2}{\sum_{j=1}^K |\mathbf{h}_k^H \mathbf{w}_j|^2 + \sigma^2} \right), \quad \forall k \in \mathcal{K}, \quad (43a)$$

$$a_k + B \log_2 \left(1 + \frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum_{j=1, j \neq k}^K |\mathbf{h}_k^H \mathbf{w}_j|^2 + \sigma^2} \right) \geq R_k, \quad \forall k \in \mathcal{K}, \quad (43b)$$

$$|\mathbf{h}_k^H \mathbf{w}_0|^2 - \sum_{j=1}^K |\mathbf{h}_k^H \mathbf{w}_j|^2 - \sigma^2 \geq \theta, \quad \forall k \in \mathcal{K}, \quad (43c)$$

$$\sum_{k=0}^K \mathbf{w}_k^H \mathbf{w}_k \leq P, \quad (43d)$$

$$a_k \geq 0, \quad \forall k \in \mathcal{K}, \quad (43e)$$

where $\mathbf{w} = [\mathbf{w}_0; \mathbf{w}_1; \dots; \mathbf{w}_K]$.

Problem (43) is hard to be solved due to nonconvex objective function and nonconvex constraints. To handle this hardness, we use the SCA approach.

B. Algorithm Design

By introducing slack variables γ_k and η_k , problem (43) can be reformulated as the following optimization problem:

$$\max_{\mathbf{a}, \mathbf{w}, \boldsymbol{\gamma}, \boldsymbol{\eta}} \sum_{k=1}^K (a_k + B \log_2 (1 + \gamma_k)), \quad (44)$$

$$\text{s.t.} \quad \sum_{k=1}^K a_k \leq B \log_2 (1 + \eta_k), \quad \forall k \in \mathcal{K}, \quad (44a)$$

$$a_k + B \log_2 (1 + \gamma_k) \geq R_k, \quad \forall k \in \mathcal{K}, \quad (44b)$$

$$\frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum_{j=1, j \neq k}^K |\mathbf{h}_k^H \mathbf{w}_j|^2 + \sigma^2} \geq \gamma_k, \quad \forall k \in \mathcal{K}, \quad (44c)$$

$$\frac{|\mathbf{h}_k^H \mathbf{w}_0|^2}{\sum_{j=1}^K |\mathbf{h}_k^H \mathbf{w}_j|^2 + \sigma^2} \geq \eta_k, \quad \forall k \in \mathcal{K}, \quad (44d)$$

$$(43c) - (43e) \quad (44e)$$

where $\boldsymbol{\gamma} = [\gamma_1, \dots, \gamma_K]^T$ and $\boldsymbol{\eta} = [\eta_1, \dots, \eta_K]^T$. For the equivalent problem (44), the objective function is concave. Problem (44) is nonconvex only due to constraints (44c), (44d), and (43c). To handle the nonconvexity of constraint (44c), we introduce new nonnegative slack variable α_k and reformulate constraints (44c) into the following equivalent constraints:

$$|\mathbf{h}_k^H \mathbf{w}_k|^2 \geq \gamma_k \alpha_k, \quad (45)$$

$$\sum_{j=1, j \neq k}^K |\mathbf{h}_k^H \mathbf{w}_j|^2 + \sigma^2 \leq \alpha_k. \quad (46)$$

Without loss of generality, the term $\mathbf{h}_k^H \mathbf{w}_k$ in constraint (45) can be expressed as a real number through an arbitrary rotation to the phase beamforming \mathbf{w}_k . As a result, constraint (45) can

be equivalent to $\mathcal{R}(\mathbf{h}_k^H \mathbf{w}_k) \geq \sqrt{\gamma_k \alpha_k}$, where $\mathcal{R}(x)$ means the real part of the complex variable x . Replacing concave function $\sqrt{\gamma_k \alpha_k}$ with the first-order Taylor series, constraint (45) becomes:

$$\mathcal{R}(\mathbf{h}_k^H \mathbf{w}_k) \geq \sqrt{\gamma_k^{(n-1)} \alpha_k^{(n-1)}} + \frac{1}{2} \sqrt{\frac{\gamma_k^{(n-1)}}{\alpha_k^{(n-1)}}} (\alpha_k - \alpha_k^{(n-1)}) + \frac{1}{2} \sqrt{\frac{\alpha_k^{(n-1)}}{\gamma_k^{(n-1)}}} (\gamma_k - \gamma_k^{(n-1)}), \quad (47)$$

where superscript $(n-1)$ means the value of the variable at the $(n-1)$ -th iteration.

Introducing new variable β_k , we can reformulate constraint (44c) as:

$$|\mathbf{h}_k^H \mathbf{w}_0|^2 \geq \beta_k \eta_k = \frac{1}{4} ((\beta_k + \eta_k)^2 - (\beta_k - \eta_k)^2), \quad (48)$$

$$\sum_{j=1}^K |\mathbf{h}_k^H \mathbf{w}_j|^2 + \sigma^2 \leq \beta_k. \quad (49)$$

Since there are K inequality constraints in (48), we cannot let the term $\mathbf{h}_k^H \mathbf{w}_0$ be a real number for all k through an arbitrary rotation to the phase beamforming \mathbf{w}_0 . To handle the nonconvexity of (48), we adopt the difference of two convex function (DC) approximation and constraint (48) can be approximated by:

$$2\mathcal{R}(\mathbf{h}_k^H \mathbf{w}_0^{(n-1)} \mathbf{h}_k^H \mathbf{w}_0) - |\mathbf{h}_k^H \mathbf{w}_0^{(n-1)}|^2 \geq \frac{1}{4} ((\beta_k + \eta_k)^2 - (\beta_k^{(n-1)} - \eta_k^{(n-1)})(\beta_k - \eta_k) + (\beta_k^{(n-1)} - \eta_k^{(n-1)})^2), \quad (50)$$

where the left hand side is the first-order Taylor series of $|\mathbf{h}_k^H \mathbf{w}_0|^2$.

By using the DC approximation, constraint (43c) can be reformulated as:

$$2\mathcal{R}(\mathbf{h}_k^H \mathbf{w}_0^{(n-1)} \mathbf{h}_k^H \mathbf{w}_0) - |\mathbf{h}_k^H \mathbf{w}_0^{(n-1)}|^2 - \sum_{j=1}^K |\mathbf{h}_k^H \mathbf{w}_j|^2 - \sigma^2 \geq \theta K \quad (51)$$

With the above approximations, the nonconvex problem (44) can be formulated in the following approximated convex problem:

$$\max_{\mathbf{a}, \mathbf{w}, \boldsymbol{\gamma}, \boldsymbol{\eta}, \boldsymbol{\alpha}, \boldsymbol{\beta}} \sum_{k=1}^K (a_k + B \log_2(1 + \gamma_k)), \quad (52)$$

$$\text{s.t.} \quad (44a), (44b), (43d), (43e), (46), (47), (49) - (51), \quad (52a)$$

$$\alpha_k \geq 0, \beta_k \geq 0, \zeta_k \geq 0, \quad \forall k \in \mathcal{K}, \quad (52b)$$

which can be effectively solved via the dual method [36].

Algorithm 2 Sum-Rate Maximization with SCA

- 1: Initialize $\mathbf{a}^{(0)}, \mathbf{w}^{(0)}, \boldsymbol{\gamma}^{(0)}, \boldsymbol{\eta}^{(0)}, \boldsymbol{\alpha}^{(0)}, \boldsymbol{\beta}^{(0)}$. Set iteration number $n = 1$.
- 2: **repeat**
- 3: Solve convex problem (52).
- 4: Denote the optimal solution of (52) by $(\mathbf{a}^{(n)}, \mathbf{w}^{(n)}, \boldsymbol{\gamma}^{(n)}, \boldsymbol{\eta}^{(n)}, \boldsymbol{\alpha}^{(n)}, \boldsymbol{\beta}^{(n)}, \boldsymbol{\zeta}^{(n)})$.
- 5: Set $n = n + 1$.
- 6: **until** the objective value (44) converges.

C. Convergence and Complexity Analysis

The SCA algorithm for solving problem (44) is summarized in Algorithm 2. Although the use of SCA and the first-order Taylor series to deal with the optimization problem is not new, we use SCA to joint optimize the beamforming and rate allocation while the works in [20]–[22] only optimized the beamforming for MISO RSMA with fixed rate allocation. The convergence of Algorithm 2 can be shown by the following lemma.

Lemma 5: The total sum-rate obtained in Algorithm 2 is monotonically non-decreasing, and the sequence $(\mathbf{a}^{(n)}, \mathbf{w}^{(n)}, \boldsymbol{\gamma}^{(n)}, \boldsymbol{\eta}^{(n)})$ converges to a point fulfilling the KKT optimal conditions of the original non-convex problem (44).

Since Lemma 5 directly follows from [37, Proposition 3], the proof of Lemma 5 is omitted. According to Algorithm 2, problem (44) is solved by using the SCA approach, where the approximated convex problem (52) is solved in each iteration. Since there are $(6K + 1)$ constraints in problem (44), the number of iterations that are required for Algorithm 2 is $\mathcal{O}(\sqrt{6K+1} \log_2(1/\epsilon))$ [38], where ϵ is the algorithm accuracy. At each iteration, the complexity of solving the second-order cone programme problem (52) is $(L_1^2 L_2)$, where $L_1 = (5 + N)K$ is the total number of variables and $L_2 = 8K + 1$ is the total number of constraints [39]. As a result, the total complexity of Algorithm 2 is $\mathcal{O}(N^2 K^{3.5} \log_2(1/\epsilon))$.

V. SIMULATION RESULTS

In this section, we evaluate the performance of the proposed optimal rate allocation and power control algorithm. There are θK users uniformly distributed in a square area of size 300 m × 300 m. The path loss model is $128.1 + 37.6 \log_{10} d$ (d is in km) [40] and the standard deviation of shadow fading is 4 dB. In addition, the bandwidth of the BS is $B = 1$ MHz and the noise power is $\sigma^2 = -104$ dBm [41]. Unless specified otherwise, the system parameters are set as maximum transmit power $P = 30$ dBm, equal rate demand $R_1 = R_2 = \dots = R_K = R = 0.5$ Mb/s and SIC detection threshold is set as $\theta = -94$ dBm¹. The main system parameters are listed in Table I. We present the results for both SISO and MISO cases. All simulation results are averaged over 100 Monte Carlo channel realizations.

¹The value of θ is selected to ensure that the error rate of decoding the common message is below a preferred requirement. A typical value of θ is set as $\theta = -94$ dBm (i.e., $\theta/\sigma^2 = 10$ dB and $\sigma^2 = -104$ dBm is the noise power) in [Table 3, 26] and [Table I, 27].

TABLE I
SYSTEM PARAMETERS

Parameter	Value
Bandwidth of the BS B	1 MHz
Noise power σ^2	-104 dBm
Maximum transmit power P	30 dBm
Minimal rate demand R	1 Mbits/s
SIC detection threshold θ	-94 dBm

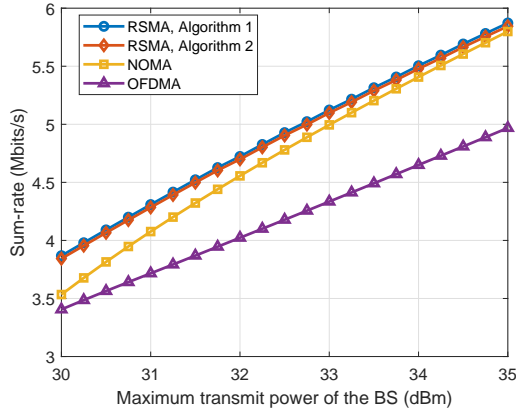


Fig. 1. Sum-rate versus maximum transmit power of the BS ($K = 3$ users, $R_1 = 1.5$ Mbits/s, $R_2 = 0.5$ Mbits/s, and $R_3 = 1$ Mbits/s).

A. SISO RSMA

In the SISO case, we do not consider the SDMA. The proposed optimal rate allocation and power control algorithm for rate maximization of RSMA is labeled as ‘RSMA’. We compare with the proposed algorithm with the optimal power control of NOMA for rate maximization [42], which is labeled as ‘NOMA’. To compare conventional OMA, we use a OFDMA system [43] as a baseline, which is labeled as ‘OFDMA’.

Fig. 1 shows the sum-rate versus maximum transmit power of the BS with unequal rate demand for three users. From this figure, we can see that RSMA always achieves the best performance among all schemes. This is due to the fact that the number of SIC in RSMA is once, while the number of SIC in NOMA can be twice, which results in high power allocation to the users with low channel gain according to the successful SIC power requirement [28, Eq. (3)] and leads to a low sum-rate. This figure also shows that RSMA outperforms NOMA particularly for low maximum BS transmit power. Compared to OFDMA, RSMA is better due to the fact that all users can be served with the whole bandwidth of the BS. According to Fig. 1, we can observe that Algorithm 1 has a similar sum-rate performance as Algorithm 2 in the SISO case. This is because both Algorithms 1 and 2 lead to suboptimal solutions to the original problem, and we set the same initial solution in these two algorithms for fair comparison. Based on the complexity analysis in Sections II-D and IV-C, we can see that the complexity of Algorithm 1 is much lower than Algorithm 2 since a closed-form solution is obtained at each step in Algorithm 1.

Fig. 2 shows the sum-rate versus minimum rate demand. From this figure, RSMA always achieves a better performance than NOMA and OFDMA. For RSMA scheme without SIC constraint, the power of the common message can be smaller than the power of the private message for some cases since

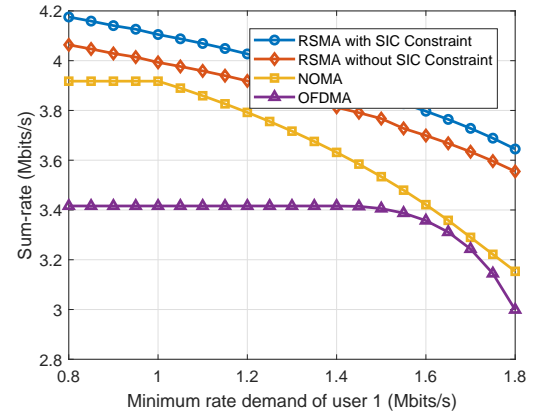


Fig. 2. Sum-rate versus minimum rate demand ($K = 3$ users, $R_2 = 0.5$ Mbits/s, and $R_3 = 1$ Mbits/s).

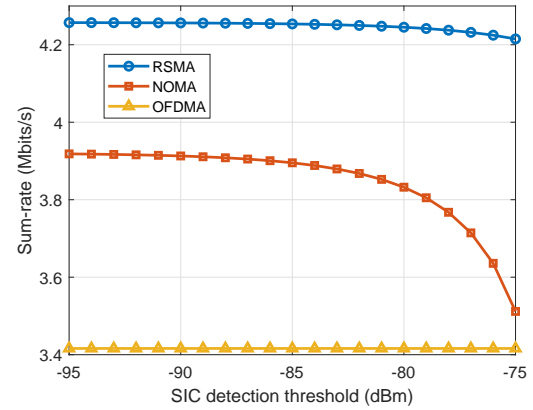


Fig. 3. Sum-rate versus SIC detection threshold ($K = 3$ users, $R_1 = 1.5$ Mbits/s, $R_2 = 0.5$ Mbits/s, and $R_3 = 1$ Mbits/s).

there is no SIC constraint. In these cases, the common message cannot be successfully decoded first and the sum-rate of RSMA scheme without SIC constraint is set as zero in the simulations. It is found that the RSMA with SIC constraint achieves a higher sum-rate than RSMA without SIC constraint. This is because some users cannot successfully decode the common message in RSMA without SIC constraint, which decreases the sum-rate. From Fig. 2, we can observe that the sum-rate decreases slightly when minimum rate demand is low. However, for a high minimum rate demand, the sum-rate decreases rapidly. This is because a high minimum rate demand requires the BS to allocate more power to the users with worse channel gains, which consequently degrades the sum-rate performance. Fig. 2 also demonstrates that, as the minimum rate demand increases, the sum-rates of OFDMA and NOMA decrease faster than RSMA. In particular, RSMA can achieve up to 15.6% and 21.5% gains in terms of data rate compared to NOMA and OFDMA, respectively. This is due to the fact that RSMA exhibits a better spectrum efficiency compared to OFDMA and NOMA, and OFDMA and NOMA are more sensitive to high minimum rate demand than RSMA.

Fig. 3 shows the sum-rate versus SIC detection threshold θ . For both RSMA and NOMA, we find that the sum-rate decreases as the SIC detection threshold increases. This is due to the fact that, as the SIC detection threshold increases, the BS must allocate more power to the common message in RSMA and the user with worse channel gain in NOMA.

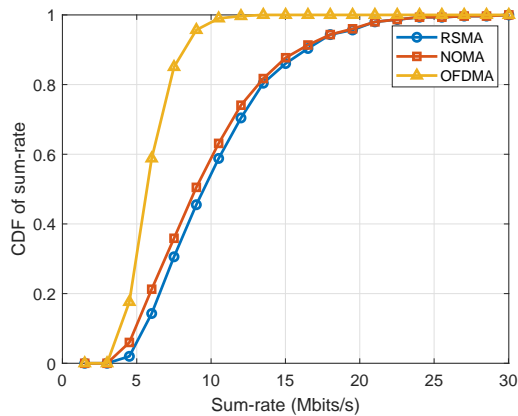


Fig. 4. CDF of the sum-rate resulting from RSMA, NOMA, and OFDMA for a network with $K = 3$ users

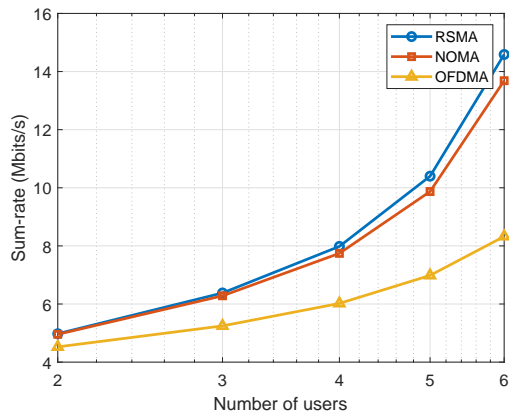


Fig. 5. Sum-rate versus number of users.

For OFDMA, naturally, the sum-rate remains the same when the SIC detection threshold increases. The proposed RSMA algorithm outperforms the NOMA in terms of sum-rate, particularly for cases with a high SIC detection threshold. Moreover, the sum-rate decreases faster for NOMA than RSMA as the SIC detection threshold increases, which implies that RSMA is more suitable for high SIC detection threshold.

Fig. 4 shows the cumulative distribution function (CDF) of the sum-rate resulting from RSMA, NOMA, and OFDMA for a network with $K = 3$ users. From Fig. 4, we observe that the CDFs for RSMA and NOMA all improve significantly compared to OFDMA, particularly for the high sum-rate region, which shows that both RSMA and NOMA are suitable for transmissions with high data rates. Moreover, we can find that RSMA outperforms NOMA at regions with moderate data rates, i.e., 5–15 Mbits/s. This is because RSMA can adjust the split between the common and private messages so as to control the interference decoding and thus optimize the sum-rate of users.

The sum-rate versus number of users is given in Fig. 5. Clearly, the proposed RSMA is always better than NOMA and OFDMA especially when the number of users is large. When the number of users is large, the multiuser gain is more apparent by the proposed RSMA compared to conventional NOMA and OFDMA. This is because RSMA can effectively determine the rate of each user receiving common message to meet its specific rate demand, while the SIC time of each user is high for NOMA and the allocated bandwidth of each

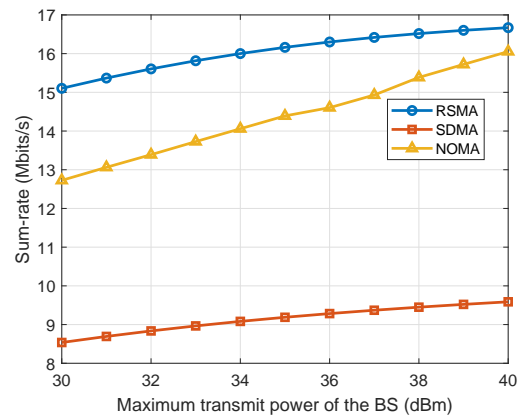


Fig. 6. Sum-rate versus maximum transmit power of the BS ($K = 3$ users and $N = 2$ antennas).

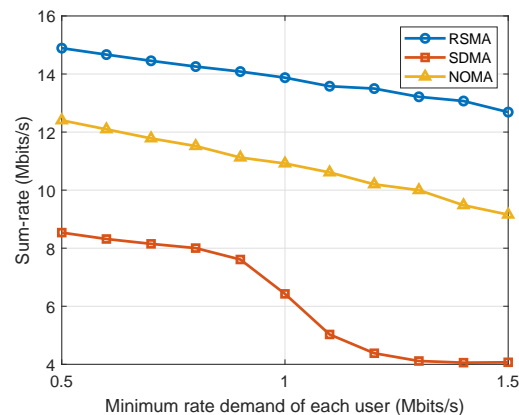


Fig. 7. Sum-rate versus minimum rate demand of each user ($K = 3$ users and $N = 2$ antennas).

user is low for OFDMA when the number of users is large. RSMA achieves better performance than NOMA and OFDMA at the cost of additional computational complexity according to Section III-D.

B. MISO RSMA

Figs. 6 and 7 respectively show how the sum-rate changes as the maximum transmit power of the BS and the minimum rate demand of each user vary for a network having three users and two antennas. From Fig. 6, we can see that the sum-rate linearly increases with the logarithmic maximum transmit power of the BS. Fig. 7 shows that the sum-rate decreases with the minimum rate demand and the decrease rate of SDMA is higher than both RSMA and NOMA. Figs. 6 and 7 show that RSMA achieves better performance than NOMA in terms of sum-rate, particularly for low maximum transmit power and high maximum rate demand. This is because RSMA dynamically allocates the rate of the common message to multiple users to meet the rate demand, while the rate decoded in NOMA is allocated to only one specific user. From Figs. 6 and 7, we observe that both RSMA and NOMA significantly outperform conventional SDMA. This is because both RSMA and NOMA use SIC to mitigate the multi-user interference, while in SDMA, each user experiences the interference from all other users.

VI. CONCLUSIONS

In this paper, we have investigated the allocation of data rate of common message transmission and the transmit power used for common and private message transmission in an RSMA system. We have formulated the problem as a sum-rate maximization problem. To solve this problem, we have derived the optimal transmit power of private message in closed form. Then, we have characterized the finite solution space for the optimal rate allocation. Simulation results show that RSMA achieves higher sum-rate than NOMA and OFDMA especially for low maximum transmit power of the BS, high minimum rate demand of users, high SIC detection threshold, and large number of users.

APPENDIX A

PROOF OF LEMMA 2

We assume that the optimal solution $(\mathbf{a}^*, \mathbf{p}^*)$ of problem (11) satisfies $\sum_{k=0}^K p_k^* < P$. We construct new power allocation by scaling power \mathbf{p}^* , i.e.,

$$p'_k = \frac{P}{\sum_{j=0}^K p_j^*} p_k^* > p_k^*, \quad \forall k = 0, 1, \dots, K. \quad (\text{A.1})$$

Given new power allocation vector $\mathbf{p}' = [p'_0, p'_1, \dots, p'_K]$, we have

$$\begin{aligned} & B \log_2 \left(1 + \frac{h_1 p'_0}{h_1 \sum_{j=1}^K p'_j + \sigma^2} \right) \\ & > B \log_2 \left(1 + \frac{h_1 p_0^*}{h_1 \sum_{j=1}^K p_j^* + \frac{\sigma^2 \sum_{j=0}^K p_j^*}{P}} \right), \end{aligned} \quad (\text{A.2})$$

$$\begin{aligned} & B \log_2 \left(1 + \frac{h_k p'_k}{h_k \sum_{j=1, j \neq k}^K p'_j + \sigma^2} \right) \\ & > B \log_2 \left(1 + \frac{h_k p_k^*}{h_k \sum_{j=1, j \neq k}^K p_j^* + \frac{\sigma^2 \sum_{j=0}^K p_j^*}{P}} \right), \end{aligned} \quad (\text{A.3})$$

and

$$p'_0 - \sum_{j=1}^K p'_j = \frac{P}{\sum_{j=0}^K p_j^*} \left(p_0^* - \sum_{j=1}^K p_j^* \right) > p_0^* - \sum_{j=1}^K p_j^* \geq \theta + \frac{\sigma^2}{h_1}, \quad (\text{A.4})$$

where the first inequalities in (A.2)-(A.4) follow from the fact that $\frac{P}{\sum_{j=0}^K p_j^*} > 1$.

According to (A.2)-(A.4), we can see that new solution $(\mathbf{a}^*, \mathbf{p}')$ is feasible and the objective value (11) of new solution is better than that of solution $(\mathbf{a}^*, \mathbf{p}^*)$, which contradicts the fact that $(\mathbf{a}^*, \mathbf{p}^*)$ is the optimal solution. Lemma 2 is proved.

APPENDIX B

PROOF OF LEMMA 3

If the pair (\mathbf{a}, \mathbf{p}) is feasible in problem (13), then the pair (\mathbf{a}, \mathbf{p}) is also feasible in problem (12) with the same objective value. It follows from the fact that the optimal value of (13) is less than or equal to the optimal value of (12).

Conversely, if the pair (\mathbf{a}, \mathbf{p}) is feasible in (12), we can construct a new pair $(\mathbf{a}', \mathbf{p})$, where

$$a'_k = \min\{a_k, R_k\}, \quad \forall k \in \mathcal{K}. \quad (\text{B.1})$$

It can be shown that solution $(\mathbf{a}', \mathbf{p})$ is feasible in problem (13). Moreover, the objective value of problem (12) is the same as problem (13). Thus, we conclude that the optimal value of (13) is greater than or equal to the optimal value of (12). Hence, problem (13) is equivalent to (12).

APPENDIX C

PROOF OF THEOREM 1

Assume that there exist m and n such that $p_m^* > p_m^{\min}$ and $p_n^* > p_n^{\min}$ for the optimal solution \mathbf{p}^* . Next, we can show that there always exist feasible power p'_m and p'_n with better objective value (14a).

To construct such p'_m and p'_n , we substitute $p_j = p_j^*, j \in \mathcal{K}, j \neq m, n$, into problem (14), which yields

$$\begin{aligned} \max_{p_m, p_n} \quad & B \log_2 \left(\frac{h_m(P - p_0) + \sigma^2}{h_m(P - p_0 - p_m) + \sigma^2} \right) \\ & + B \log_2 \left(\frac{h_n(P - p_0) + \sigma^2}{h_n(P - p_0 - p_n) + \sigma^2} \right), \end{aligned} \quad (\text{C.1})$$

$$\text{s.t.} \quad p_m + p_n = P - p_0 - \sum_{j=1, j \neq m, n}^K p_j^*, \quad (\text{C.1a})$$

$$p_m \geq p_m^{\min}, p_n \geq p_n^{\min}. \quad (\text{C.1b})$$

According to (C.1a), we have

$$p_m = P - p_0 - \sum_{j=1, j \neq m, n}^K p_j^* - p_n. \quad (\text{C.2})$$

Combining (C.1b) and (C.2), we have

$$p_n^{\min} \leq p_n \leq P - p_0 - \sum_{j=1, j \neq m, n}^K p_j^* - p_m^{\min}. \quad (\text{C.3})$$

From (C.2) and (C.3), we can see that problem (C.1) can be simplified as

$$\begin{aligned} \max_{p_n} \quad & -B \log_2 \left(h_m \left(\sum_{j=1, j \neq m, n}^K p_j^* + p_n \right) + \sigma^2 \right) \\ & -B \log_2 (h_n(P - p_0 - p_n) + \sigma^2), \end{aligned} \quad (\text{C.4})$$

$$\text{s.t.} \quad p_n^{\min} \leq p_n \leq P - p_0 - \sum_{j=1, j \neq m, n}^K p_j^* - p_m^{\min}. \quad (\text{C.4a})$$

Due to the convexity of function $-\log(x)$, the objective function (C.4) is convex. Since the maximization of a convex function always lies in the boundary of the feasible solution, i.e., the optimal solution p'_n of problem (C.4) satisfies

$$p'_n \in \left\{ p_n^{\min}, P - p_0 - \sum_{j=1, j \neq m, n}^K p_j^* - p_m^{\min} \right\}. \quad (\text{C.5})$$

Further considering (C.2), we can construct

$$p'_m = P - p_0 - \sum_{j=1, j \neq m, n}^K p_j^* - p'_n. \quad (\text{C.6})$$

Based on the equivalence of problem (C.1) and problem (C.4), (p'_m, p'_n) is the optimal solution of problem (C.1). According to (C.5) and (C.6), $p'_m = p_m^{\min}$ or $p'_n = p_n^{\min}$ is always satisfied.

Since (p'_m, p'_n) is the optimal solution of problem (C.1) and $(p'_m, p'_n) \neq (p_m^*, p_n^*)$, we can claim that solution

$$(p_1^*, \dots, p_{m-1}^*, p'_m, p_{m+1}^*, \dots, p_{n-1}^*, p'_n, p_{n+1}^*, \dots, p_K^*) \quad (\text{C.7})$$

is feasible with better objective value than solution \bar{p}^* , which contradicts the fact that \bar{p}^* is the optimal solution of problem (14).

As a result, the proof of Theorem 1 is complete.

APPENDIX D PROOF OF THEOREM 2

Substituting the optimal power allocation $p_k^* = P - p_0 - \sum_{j=1, j \neq k}^K p_j^{\min}$ and $p_j^* = p_j^{\min}, \forall j \in \mathcal{K}, j \neq k$ to problem (14) according to Theorem 1, we can obtain the objective value (14a) as

$$\begin{aligned} & \sum_{j=1, j \neq k}^K B \log_2 \left(\frac{h_j(P - p_0) + \sigma^2}{h_j(P - p_0 - p_j^{\min}) + \sigma^2} \right) \\ & + B \log_2 \left(\frac{h_k(P - p_0) + \sigma^2}{h_k \sum_{j=1, j \neq k}^K p_j^{\min} + \sigma^2} \right). \end{aligned} \quad (\text{D.1})$$

To maximize sum-rate (D.1), the optimal k should be chosen as

$$\begin{aligned} k &= \arg \max_{m \in \mathcal{K}} \sum_{j=1}^K B \log_2 \left(\frac{h_j(P - p_0) + \sigma^2}{h_j(P - p_0 - p_j^{\min}) + \sigma^2} \right) \\ & - B \log_2 \left(\frac{h_m(P - p_0) + \sigma^2}{h_m(P - p_0 - p_m^{\min}) + \sigma^2} \right) \\ & + B \log_2 \left(\frac{h_m(P - p_0) + \sigma^2}{h_m \sum_{j=1, j \neq m}^K p_j^{\min} + \sigma^2} \right) \\ & = \arg \max_{m \in \mathcal{K}} B \log_2 (h_m(P - p_0 - p_m^{\min}) + \sigma^2) \\ & - B \log_2 \left(h_m \sum_{j=1, j \neq m}^K p_j^{\min} + \sigma^2 \right) \\ & = \arg \max_{m \in \mathcal{K}} \frac{h_m(P - p_0 - p_m^{\min}) + \sigma^2}{h_m \sum_{j=1, j \neq m}^K p_j^{\min} + \sigma^2} - 1 \\ & = \arg \max_{m \in \mathcal{K}} \frac{P - p_0 - \sum_{j=1}^K p_j^{\min}}{\sum_{j=1}^K p_j^{\min} + \frac{\sigma^2}{h_m} - p_m^{\min}} \\ & = \arg \min_{m \in \mathcal{K}} \frac{\sigma^2}{h_m} - p_m^{\min}. \end{aligned} \quad (\text{D.2})$$

Substituting (15) to (D.2), we have

$$\begin{aligned} k &= \arg \min_{j \in \mathcal{K}} \frac{\sigma^2}{h_j} - \left(1 - 2^{\frac{a_j - R_j}{B}}\right) \left(P - p_0 + \frac{\sigma^2}{h_j}\right) \\ & = \arg \min_{j \in \mathcal{K}} 2^{\frac{a_j - R_j}{B}} \left(P - p_0 + \frac{\sigma^2}{h_j}\right). \end{aligned} \quad (\text{D.3})$$

This completes the proof of Theorem 2.

APPENDIX E

PROOF OF THEOREM 3

We first show that $a_{\pi_m}^* \geq a_{\pi_n}^*$ for all $m < n$ for the optimal solution \mathbf{a}^* of problem (28). This can be proved by the contradiction method. If there exists $m < n$ such that $a_{\pi_m}^* < a_{\pi_n}^*$, we can construct a new solution \mathbf{a}' with $a'_{\pi_m} = a_{\pi_n}^*$, $a'_{\pi_n} = a_{\pi_m}^*$, $a'_j = a_j^*$ for $j \neq m, n$. Then, we have:

$$\begin{aligned} & \sum_{j=1}^K \left(1 - 2^{\frac{a'_{\pi_j} - R_{\pi_j}}{B}}\right) \left(P - p_0 + \frac{\sigma^2}{h_{\pi_j}}\right) \\ & = \sum_{j=1}^K \left(1 - 2^{\frac{a_{\pi_j}^* - R_{\pi_j}}{B}}\right) \left(P - p_0 + \frac{\sigma^2}{h_{\pi_j}}\right) \\ & \quad + \left(2^{\frac{a_m^*}{B}} - 2^{\frac{a_n^*}{B}}\right) \left(2^{\frac{-R_{\pi_m}}{B}} \left(P - p_0 + \frac{\sigma^2}{h_{\pi_m}}\right) - 2^{\frac{-R_{\pi_n}}{B}} \left(P - p_0 + \frac{\sigma^2}{h_{\pi_n}}\right)\right) \\ & \leq \sum_{j=1}^K \left(1 - 2^{\frac{a_{\pi_j}^* - R_{\pi_j}}{B}}\right) \left(P - p_0 + \frac{\sigma^2}{h_{\pi_j}}\right), \end{aligned} \quad (\text{E.1})$$

where the inequality follows from (29). Based on (E.1), we can claim that the new solution \mathbf{a}' is feasible with better or equal objective value than solution \mathbf{a}^* .

Then, we show that the objective function (28) monotonically increases with a_j for $j \neq k$. To show this, the first derivative of the objective function (28) with respect to a_j can be given by:

$$\begin{aligned} \frac{\partial f(\mathbf{a})}{\partial a_j} &= -1 + \frac{2^{\frac{a_j - R_j}{B}} \left(P - p_0 + \frac{\sigma^2}{h_j}\right)}{\sum_{j=1, j \neq k}^K \left(1 - 2^{\frac{a_j - R_j}{B}}\right) \left(P - p_0 + \frac{\sigma^2}{h_j}\right) + \frac{\sigma^2}{h_k}} \\ & \geq -1 + \frac{2^{\frac{a_k - R_k}{B}} \left(P - p_0 + \frac{\sigma^2}{h_k}\right)}{\sum_{j=1, j \neq k}^K \left(1 - 2^{\frac{a_j - R_j}{B}}\right) \left(P - p_0 + \frac{\sigma^2}{h_j}\right) + \frac{\sigma^2}{h_k}} \\ & = \frac{P - p_0 - \sum_{j=1}^K \left(1 - 2^{\frac{a_j - R_j}{B}}\right) \left(P - p_0 + \frac{\sigma^2}{h_j}\right) - \frac{\sigma^2}{h_k}}{\sum_{j=1, j \neq k}^K \left(1 - 2^{\frac{a_j - R_j}{B}}\right) \left(P - p_0 + \frac{\sigma^2}{h_j}\right) + \frac{\sigma^2}{h_k}} \geq 0, \end{aligned} \quad (\text{E.2})$$

where $f(\mathbf{a})$ denotes the objective function (28). The first inequality in (E.2) follows from constraint (28b), and the second inequality in (E.2) follows from constraint (28c).

Based on (E.2), the objective function (28) increases with each rate a_j , while (E.1) shows that it is optimal to allocate more rate to the user that has the lower channel gain. As a result, the optimal rate allocation can be given in (30) and (31). This completes the proof of Theorem 3.

APPENDIX F

PROOF OF THEOREM 4

For equal data rate requirement $R_1 = \dots = R_K = R$, constraint (29) becomes:

$$2^{\frac{-R}{B}} \left(P - p_0 + \frac{\sigma^2}{h_{\pi_j}}\right) \geq 2^{\frac{-R}{B}} \left(P - p_0 + \frac{\sigma^2}{h_{\pi_l}}\right), \quad \forall j < l, \quad (\text{F.1})$$

i.e., $h_{\pi_j} \leq h_{\pi_l}$ for all $j < l$. Since $h_1 \leq \dots \leq h_K$, we have $\pi_j = j$.

If $c_1 \in [(l-1)R, lR)$ with $0 \leq l < K$, i.e.,

$$p_0 \in \left[\left(1 - 2^{-(l-1)R}\right) \left(P + \frac{\sigma^2}{h_1}\right), \left(1 - 2^{-lR}\right) \left(P + \frac{\sigma^2}{h_1}\right) \right]. \quad (\text{F.2})$$

According to Theorem 3, the optimal rate allocation is:

$$a_j^* = \begin{cases} R, & \text{if } j < l, \\ c_1 - (l-1)R, & \text{if } j = l, \\ 0, & \text{otherwise.} \end{cases} \quad (\text{F.3})$$

Substituting (F.3) and (32) into problem (27) yields:

$$\max_{p_0} B \log_2 \left(\frac{P - p_0 + \frac{\sigma^2}{h_k}}{\sum_{j=1}^{K-1} \left(1 - 2^{\frac{a_j^* - R}{B}}\right) \left(P - p_0 + \frac{\sigma^2}{h_j}\right) + \frac{\sigma^2}{h_k}} \right), \quad (\text{F.4})$$

$$\text{s.t.} \quad \left(1 - \frac{h_1 P + \sigma^2}{h_1(P - p_0) + \sigma^2} 2^{\frac{-R}{B}}\right) \left(P - p_0 + \frac{\sigma^2}{h_j}\right) + \sum_{j=l}^K \left(1 - 2^{\frac{-R}{B}}\right) \left(P - p_0 + \frac{\sigma^2}{h_j}\right) \leq P - p_0, \quad (\text{F.4a})$$

$$p_0 \geq \frac{P}{2} + \frac{\theta + \sigma^2}{2h_1}, \quad (\text{F.4b})$$

$$p_0 \in \left[\left(1 - 2^{-KR}\right) \left(P + \frac{\sigma^2}{h_1}\right), \left(1 - 2^{-(l-1)R}\right) \left(P + \frac{\sigma^2}{h_1}\right) \right]. \quad (\text{F.4c})$$

Through calculating the first-order derivative of objective function (F.4), the objective function (F.4) is always monotonically increasing/decreasing with any given a_j^* . As a result, the optimal p_0^* always lies in the corner points of feasible set (F.4a)-(F.4c), as shown in Theorem 4.

If $c_1 \geq KR$, we have $p_0 \geq (1 - 2^{-KR}) \left(P + \frac{\sigma^2}{h_1}\right)$. The optimal rate allocation is $a_j^* = R$, $\forall j \in \mathcal{K}$, according to Theorem 3. Applying $a_j^* = R$ to problem (27) yields:

$$\max_{p_0} B \log_2 \left(\frac{P - p_0 + \frac{\sigma^2}{h_k}}{\frac{\sigma^2}{h_k}} \right), \quad (\text{F.5})$$

$$\text{s.t.} \quad p_0 \geq \frac{P}{2} + \frac{\theta + \sigma^2}{2h_1}, \quad (\text{F.5a})$$

$$p_0 \geq (1 - 2^{-KR}) \left(P + \frac{\sigma^2}{h_1}\right). \quad (\text{F.5b})$$

Since the objective function (F.5) monotonically decreases with p_0 , the optimal p_0^* lies in case 1 of Theorem 4. This completes the proof of Theorem 4.

APPENDIX G PROOF OF THEOREM 5

According to Lemma 1, problem (27) with given rate allocation \mathbf{a} becomes:

$$\max_{p_0} B \log_2 \left(\frac{P - p_0 + \frac{\sigma^2}{h_k}}{\sum_{j=1, j \neq k}^K \left(1 - 2^{\frac{a_j - R}{B}}\right) \left(P - p_0 + \frac{\sigma^2}{h_j}\right) + \frac{\sigma^2}{h_k}} \right), \quad (\text{G.1})$$

$$\text{s.t.} \quad \sum_{j=1}^K a_j \leq B \log_2 \left(\frac{h_1 P + \sigma^2}{h_1(P - p_0) + \sigma^2} \right), \quad (\text{G.1a})$$

$$\sum_{j=1}^K \left(1 - 2^{\frac{a_j - R}{B}}\right) \left(P - p_0 + \frac{\sigma^2}{h_j}\right) \leq P - p_0, \quad (\text{G.1b})$$

$$p_0 \geq \frac{P}{2} + \frac{\theta + \sigma^2}{2h_1}. \quad (\text{G.1c})$$

Since the objective function (G.1) is always monotonically increasing/decreasing, the optimal p_0^* always lies in the corner points of feasible set (G.1a)-(G.1c), as shown in Theorem 5.

REFERENCES

- [1] Z. Yang, M. Chen, W. Saad, and M. Shikh-Bahaei, "Downlink sum-rate maximization for rate splitting multiple access (RSMA)," in *Proc. IEEE Int. Conf. Commun.*, Dublin, Ireland, June 2020.
- [2] C. Chaccour, M. N. Soorki, W. Saad, M. Bennis, and P. Popovski, "Can terahertz provide high-rate reliable low latency communications for wireless VR?" *arXiv preprint arXiv:2005.00536*, 2020.
- [3] F. Hu, Y. Deng, W. Saad, M. Bennis, and A. H. Aghvami, "Cellular-connected wireless virtual reality: Requirements, challenges, and solutions," *IEEE Commun. Mag.*, vol. 58, no. 5, pp. 105–111, May 2020.
- [4] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, to appear, 2020.
- [5] Y. Liu, Z. Qin, M. ElKashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Nonorthogonal multiple access for 5G and beyond," *IEEE Proceedings*, vol. 105, no. 12, pp. 2347–2381, Dec. 2017.
- [6] L. Dai, B. Wang, Y. Yuan, S. Han, C. I. I, and Z. Wang, "Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sep. 2015.
- [7] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. K. Bhargava, "A survey on non-orthogonal multiple access for 5G networks: Research challenges and future trends," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2181–2195, Oct. 2017.
- [8] Z. Yang, C. Pan, W. Xu, Y. Pan, M. Chen, and M. ElKashlan, "Power control for multi-cell networks with non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 927–942, Feb. 2018.
- [9] L. Yang, Q. Ni, L. Lv, J. Chen, X. Xue, H. Zhang, and H. Jiang, "Cooperative non-orthogonal layered multicast multiple access for heterogeneous networks," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1148–1165, Feb. 2019.
- [10] Y. Li, H. Zhang, K. Long, C. Jiang, and M. Guizani, "Joint resource allocation and trajectory optimization with QoS in UAV-based NOMA wireless networks," *IEEE Trans. Wireless Commun.*, pp. 1–1, 2021.
- [11] K. Wang, Y. Liu, Z. Ding, and A. Nallanathan, "User association in non-orthogonal multiple access networks," in *Proc. IEEE Int. Conf. Commun.*, Kansas City, MO, USA, May 2018, pp. 1–6.
- [12] 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 Veh. Technol. Conf.* Dresden, German, June 2013, pp. 1–5.
- [13] Z. Yang, W. Xu, C. Pan, Y. Pan, and M. Chen, "On the optimality of power allocation for NOMA downlinks with individual QoS constraints," *IEEE Commun. Lett.*, vol. 21, no. 7, pp. 1649–1652, July 2017.
- [14] B. Rimoldi and R. Urbanke, "A rate-splitting approach to the gaussian multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 364–375, Mar. 1996.
- [15] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.
- [16] B. Clerckx, H. Joudeh, C. Hao, M. Dai, and B. Rassouli, "Rate splitting for MIMO wireless networks: A promising PHY-layer strategy for LTE evolution," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 98–105, May 2016.

- [17] J. Cao and E. M. Yeh, "Asymptotically optimal multiple-access communication via distributed rate splitting," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 304–319, Jan. 2007.
- [18] Y. Mao, B. Clerckx, and V. O. K. Li, "Rate-splitting multiple access for downlink communication systems: Bridging, generalizing, and outperforming SDMA and NOMA," *EURASIP J. Wireless Commun. Network.*, vol. 2018, no. 1, pp. 1–54, May 2018.
- [19] —, "Rate-splitting for multi-user multi-antenna wireless information and power transfer," *arXiv preprint arXiv:1902.07851*, 2019.
- [20] H. Joudeh and B. Clerckx, "Sum-rate maximization for linearly precoded downlink multiuser MISO systems with partial CSIT: A rate-splitting approach," *IEEE Trans. Commun.*, vol. 64, no. 11, pp. 4847–4861, Nov. 2016.
- [21] Y. Mao, B. Clerckx, and V. O. K. Li, "Energy efficiency of rate-splitting multiple access, and performance benefits over SDMA and NOMA," in *Proc. IEEE Int. Symp. Wireless Commun. Sys.*, Lisbon, Portugal, Aug. 2018, pp. 1–5.
- [22] M. Dai and B. Clerckx, "Multiuser millimeter wave beamforming strategies with quantized and statistical CSIT," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7025–7038, Nov. 2017.
- [23] A. Rahmati, Y. Yapıcı, N. Rupasinghe, I. Guvenc, H. Dai, and A. B. Huyany, "Energy efficiency of RSMA and NOMA in cellular-connected mmwave UAV networks," *arXiv preprint arXiv:1902.04721*, Feb. 2019.
- [24] H. Joudeh and B. Clerckx, "Robust transmission in downlink multiuser MISO systems: A rate-splitting approach," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6227–6242, Dec. 2016.
- [25] C. Hao, Y. Wu, and B. Clerckx, "Rate analysis of two-receiver MISO broadcast channel with finite rate feedback: A rate-splitting approach," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3232–3246, Sep. 2015.
- [26] Y. Mao, B. Clerckx, and V. O. K. Li, "Rate-splitting for multi-antenna non-orthogonal unicast and multicast transmission: Spectral and energy efficiency analysis," *IEEE Trans. Commun.*, to appear, 2019.
- [27] M. S. Ali, H. Tabassum, and E. Hossain, "Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (NOMA) systems," *IEEE ACCESS*, vol. 4, pp. 6325–6343, Aug. 2016.
- [28] M. S. Ali, E. Hossain, A. Al-Dweik, and D. I. Kim, "Downlink power allocation for CoMP-NOMA in multi-cell networks," *IEEE Trans. Commun.*, vol. 66, no. 9, pp. 3982–3998, Sep. 2018.
- [29] M. F. Hanif, Z. Ding, T. Ratnarajah, and G. K. Karagiannidis, "A minorization-maximization method for optimizing sum rate in the downlink of non-orthogonal multiple access systems," *IEEE Trans. Signal Process.*, vol. 64, no. 1, pp. 76–88, 2015.
- [30] X. Chen, Z. Zhang, C. Zhong, and D. W. K. Ng, "Exploiting multiple-antenna techniques for non-orthogonal multiple access," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2207–2220, Oct. 2017.
- [31] S. R. Islam, N. Avazov, O. A. Dobre, and K.-S. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges," *IEEE Commun. Surveys Tut.*, vol. 19, no. 2, pp. 721–742, 2016.
- [32] B. Clerckx, Y. Mao, R. Schober, and H. V. Poor, "Rate-splitting unifying SDMA, OMA, NOMA, and multicasting in MISO broadcast channel: A simple two-user rate analysis," *arXiv preprint arXiv:1906.04474*, 2019.
- [33] Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, "Optimal joint power and subcarrier allocation for full-duplex multicarrier non-orthogonal multiple access systems," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1077–1091, Mar. 2017.
- [34] H. Joudeh and B. Clerckx, "Rate-splitting for max-min fair multigroup multicast beamforming in overloaded systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7276–7289, Nov. 2017.
- [35] A. Papazafeiropoulos, B. Clerckx, and T. Ratnarajah, "Rate-splitting to mitigate residual transceiver hardware impairments in massive mimo systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 9, pp. 8196–8211, Sept. 2017.
- [36] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.
- [37] A. Zappone, E. Bjornson, L. Sanguinetti, and E. Jorswieck, "Globally optimal energy-efficient power control and receiver design in wireless networks," *IEEE Trans Signal Process.*, vol. 65, no. 11, pp. 2844–2859, June 2017.
- [38] M. Grant, S. Boyd, and Y. Ye, "Cvx: Matlab software for disciplined convex programming," 2008.
- [39] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear algebra and its applications*, vol. 284, no. 1-3, pp. 193–228, 1998.
- [40] M. Chen, Z. Yang, W. Saad, C. Yin, S. Cui, and H. V. Poor, "A joint learning and communications framework for federated learning over wireless networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 269–283, Jan. 2021.
- [41] M. Mozaffari, A. Taleb Zadeh Kasgari, W. Saad, M. Bennis, and M. Debbah, "Beyond 5G with UAVs: Foundations of a 3D wireless cellular network," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 357–372, Jan. 2019.
- [42] J. Zhu, J. Wang, Y. Huang, S. He, X. You, and L. Yang, "On optimal power allocation for downlink non-orthogonal multiple access systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2744–2757, 2017.
- [43] K. Seong, M. Mohseni, and J. M. Cioffi, "Optimal resource allocation for OFDMA downlink systems," in *Proc. IEEE Int. Symp. Inf. Theory*, Seattle, WA, USA, 2006, pp. 1394–1398.