This article has been accepted for publication in IEEE Transactions on Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2025.3616144

Optimal Power Allocation for Multi-Group Non-Orthogonal Multiple Access Systems

Qian Peng, *Student Member, IEEE*, Rui Zhou, *Member, IEEE*, Jiaheng Wang, *Senior Member, IEEE*, Xiang-Gen Xia, *Fellow, IEEE*, Zhiguo Ding, *Fellow, IEEE*

***Abstract*—Non-orthogonal multiple access (NOMA) is recognized as a pivotal technology for enhancing spectral efficiency and facilitating massive connectivity in wireless communications. In practice, users are often divided into groups that occupy different resource blocks, while the users in the same group share the same resource through power-domain multiplexing, leading to multi-group NOMA (MG-NOMA). Hence, the full benefit of MG-NOMA relies on optimum power allocation, which, unfortunately, leads to difficult optimization problems that have not been well solved so far. This work investigates the optimal power allocation to achieve the maximum spectral efficiency and energy efficiency with quality-of-service (QoS) requirements and arbitrary user weights in MG-NOMA. We show that the complicated MG-NOMA power allocation problems can be decomposed into two layers, the lower layer for power allocation within each group and the upper layer for the power budget allocation among groups. More importantly, we reveal that, although quite difficult, the two-layer problems both contain the hidden convexity, indicating the achievability of the globally optimal solution. Then, we derive the optimal power allocation in either closed or semi-closed form for the maximum spectral efficiency and energy efficiency in MG-NOMA, respectively. The simulation results demonstrate the superiority of the proposed optimal power allocation schemes.**

***Index Terms*—Non-orthogonal multiple access, weighted sum rate maximization, weighted energy efficiency maximization, quality of service, closed-form.**

I. Introduction

The rapid advancement of internet of things (IoT) applications is propelling the development of sixth-generation (6G) communication networks, which is expected to accommodate substantial data traffic and device connectivity while managing resource limitations [1]–[4]. To meet these challenges, there is a critical need for innovative multiple access strategies that efficiently utilize resources and manage system complexity. Unlike traditional orthogonal multiple access (OMA) methods such as time division multiple access (TDMA) and frequency division multiple access (FDMA), which are limited by their capacity to support only a finite number of users due to the orthogonal usage of resources, non-orthogonal multiple access (NOMA), where multiple users are superposed in the same time and frequency resources, offers a more efficient solution with additional multiplexing and significantly enhances multiple access efficiency [5]–[8].

Initial studies on NOMA focused on single-group NOMA (SG-NOMA), wherein different users are allocated different power levels in the same resource block, such as a time slot or carrier. The user power difference is exploited to detect the desired signal via successive interference cancellation (SIC) at the receiver side. Thus, the effectiveness of SG-NOMA critically depends on proper power allocation. Various power allocation strategies have been explored for SG-NOMA, such as fairness-based optimization [9], total power minimization [10], minimum error probability designs [11], [12], and sum rate maximization [13], [14]. Energy-efficient power allocation has been investigated in [15]–[17], and outage performance has been evaluated in [18]. However, the served user number in a single group is limited as the users in SG-NOMA systems suffer poor communication performance with a larger number of users. Therefore, to enhance system performance, excessive users are generally divided into multiple groups in practical scenarios.

Consequently, research attention has pivoted to multi-group NOMA (MG-NOMA) [19], where the users are divided into several groups. Particularly, the users in one group are assigned to the same resource block following the SG-NOMA principle and different user groups are assigned to different orthogonal resource blocks, which effectively avoids inter-group interference. Hence, MG-NOMA is more flexible to further improve spectral efficiency and facilitate massive connectivity [20]. Yet, power allocation in MG-NOMA systems becomes more complex since it requires to consider both intra-group and inter-group power allocation. A heuristic power allocation design [21] and several learning-based methods [22]–[26] have been proposed for MG-NOMA, in which only the suboptimal solutions are provided. Zhu et al., in [27], first derived the optimal power allocation for MG-NOMA under various performance criteria, but the served user number per group was restricted to two. In line with this work, several researchers explored the power optimization strategies for MG-NOMA with multiple users per group. Particularly, the work [28] investigated the weighted sum rate maximization problem in MG-NOMA systems and tried to obtain the optimal power

This work was supported in part by the Natural Science Foundation on Frontier Leading Technology Basic Research Project of Jiangsu under Grants BK20212001 and BK20222001, the National Natural Science Foundation of China under Grants 62331024, U22B2006 and 62201362, the Key Technologies R&D Program of Jiangsu (Prospective and Key Technologies for Industry) under Grant BE2022068-3, The Taihu Lake Innovation Fund for the School of Future Technology of Southeast University, the Fundamental Research Funds for the Central Universities under Grants 2242022K60002 and 2242022K60001, the Jiangsu Provincial Scientific Research Center of Applied Mathematics under Grant BK20233002, the Southeast University-China Mobile Research Institute Joint Innovation Center. Corresponding author: Jiaheng Wang.

Qian Peng and Jiaheng Wang are with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing 210096, China, and also with Purple Mountain Laboratories, Nanjing 211111, China (e-mail: {qianpeng, jhwang}@seu.edu.cn). Ruizhou is with Shenzhen Research Institute of Big Data, The Chinese University of Hong Kong-Shenzhen, Shenzhen 518172, China (e-mail: rui.zhou@sribd.cn). Xiang-Gen Xia is with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716, USA (e-mail: xxia@ee.udel.edu). Zhiguo Ding is with the School of Electrical and Electronic Engineering, Manchester University, Manchester M13 9PL, U.K. (e-mail: zhiguo.ding@manchester.ac.uk).

allocation via exhaustive search, which leads to a forbidden computational complexity. Later, the work [29] obtained the optimal power allocation for weighted sum rate maximization in MG-NOMA systems. However, in [29] the user weights are confined such that higher weights were designated to users with inferior channel conditions. Consequently, so far, the optimal power allocation for MG-NOMA with a general performance metric, e.g., sum rate or energy efficiency maximization with any weights and quality-of-service (QoS) requirements, is still unknown.

In this paper, we consider, from a most general perspective, MG-NOMA systems with an arbitrary number of users divided into an arbitrary number of groups, where each group is served via a single orthogonal resource block. We consider two most common performance metrics, the weighted sum rate (WSR) and weighted energy efficiency (WEE) of users, where no restriction is imposed on the user weights or power orders. We also consider QoS constraints for weighted sum rate maximization (WSRM) and weighted energy efficiency maximization (WEEM) to guarantee individual performance for each user. We show that the optimal MG-NOMA power allocation schemes, for WSRM and WEEM under QoS constraints, can be analytically derived or even obtained in closed-form along with important insights. The key contributions of our work are outlined as follows:

- We consider a general MG-NOMA system, and formulate two power allocation problems, i.e., the WSRM and WEEM problems, both under user QoS constraint, without any restriction on weights or power orders.
- We show that the formulated MG-NOMA power allocation problem can be equivalently decomposed into two layers, the upper layer for power budget allocation among groups and the lower layer for power allocation within each group.
- For the SG-NOMA power allocation problem, we prove that, without loss of optimality, a group can be divided into several independent subgroups. Each subgroup can be treated as a virtual user and has an independent power allocation strategy.
- Based on this finding, we are able to derive the optimal intra-group power allocation in closed-form for SG-NOMA. We further analytically characterize the maximum WSR of each group with respect to its power budget. Then, we prove that the maximum WSR for each group is convex in the power budgets allocated to the groups, and finally analytically characterize the globally optimal solution to the MG-NOMA WSRM problem.
- For the MG-NOMA WEEM problem, we first prove that WEEM is equivalent to WSRM in some cases and then provide a simple method to obtain the optimal power allocation in the general case. We further provide a closed-form solution to the MG-NOMA WEEM problem under some mild conditions.
- Numerical results are provided to demonstrate that our proposed power allocation strategies outperform the existing schemes and achieve optimal performance with low computational complexity.

The rest of the paper is organized as follows. The system model and problem formulations are outlined in Section II. We investigate the optimal power allocation for WSRM in SG-NOMA systems and MG-NOMA systems in Section III. The optimal power allocation for the WEEM problem in MG-NOMA systems is presented in Section IV. The numerical results are presented to evaluate the performance of the proposed power allocation schemes in Section V, and finally, conclusions are drawn in Section VI.

II. System Model and Problem Formulation

A. System Model

Consider the downlink of a cellular MG-NOMA system wherein a base station (BS) serves N users through M resource blocks, i.e., M orthogonal carriers, indexed by $m = 1, 2, \dots, M$. The total bandwidth B is divided into M groups, so the bandwidth for each group is $B_c = B/M$. The BS and users are equipped with a single antenna.[1] Let N_m be the number of users within group m for $m = 1, 2, \dots, M$, and $\mathrm{UE}_{n,m}$ denotes user n of group m for $m = 1, 2, \dots, M$. In practice, each user is often intended to use only one group, thus we focus on this typical situation in this paper and have $N = \sum_{m=1}^{M} N_m$. The signal transmitted by the BS of group m can be formulated as

$$x_m = \sum_{n=1}^{N_m} \sqrt{p_{n,m}} s_{n,m},$$

where $s_{n,m}$ represents the message intended for $\mathrm{UE}_{n,m}$ with $\mathbb{E}[|s_{n,m}|^2] = 1$, and $p_{n,m}$ represents the power allocated to $\mathrm{UE}_{n,m}$ by the BS for sending $s_{n,m}$. The received signal at $\mathrm{UE}_{n,m}$ can be formulated as

$$y_{n,m} = h_{n,m} \sum_{i=1}^{N_m} \sqrt{p_{i,m}} s_{i,m} + z_{n,m},$$

where $h_{n,m}$ represents the channel coefficient from the BS to $\mathrm{UE}_{n,m}$, and $z_{n,m}$ is the complex additive white Gaussian noise (AWGN) with zero-mean and variance $\sigma_{n,m}^2$, i.e., $z_{n,m} \sim \mathcal{CN}(0, \sigma_{n,m}^2)$. Let $\alpha_{n,m} \triangleq |h_{n,m}|^2/\sigma_{n,m}^2$ and assume without loss of generality that the users are ordered by their normalized channel gains such that $\alpha_{1,m} \leq \alpha_{2,m} \leq \cdots \leq \alpha_{N_m,m}$. Then, it is often assumed that the perfect channel state information (CSI) is available at both the BS and each user. The SIC receiver at $\mathrm{UE}_{n,m}$ eliminates the influence of weaker users, i.e., $i < n$ [11], [13], [28], [30]–[32]. This is an assumption that applies to all users. Thus, the signal to interference-plus-noise ratio (SINR) of $\mathrm{UE}_{n,m}$ is

$$\mathrm{SINR}_{n,m} = \frac{|h_{n,m}|^2 p_{n,m}}{|h_{n,m}|^2 \sum_{i>n}^{N_m} p_{i,m} + \sigma_{n,m}^2}.$$

Hence, the data rate of $\mathrm{UE}_{n,m}$ can be formulated as

[1]Note that although the results obtained in this work are based on the single-antenna scenarios, they can also provide important insights and be extended to NOMA systems with multiple antennas.

$$\begin{aligned} R_{n,m} &\triangleq B_c \log\left(1+\frac{|h_{n,m}|^2 p_{n,m}}{|h_{n,m}|^2\sum_{i>n}^{N_m} p_{i,m}+\sigma_{n,m}^2}\right) \\ &= B_c \log\left(1+\frac{\alpha_{n,m}p_{n,m}}{\alpha_{n,m}\sum_{i>n}^{N_m} p_{i,m}+1}\right). \end{aligned} \tag{1}$$

B. Problem Formulation

In this paper, we focus on the power allocation strategy in MG-NOMA systems. The most common objective is the weighted sum rate of all users. User weights are introduced to strike a balance between spectral efficiency and user fairness. To avoid no power allocation for the users, the QoS constraints shall also be considered. Therefore, we examine the WSRM problem:

$$\begin{aligned} \mathcal{P}_1^{\rm WSRM}: \underset{\{\mathbf{p}_m\}_{m=1}^M}{\text{maximize}} \ & \sum_{m=1}^{M}\sum_{n=1}^{N_m} w_{n,m}R_{n,m} \\ \text{subject to} \ & \sum_{m=1}^{M}\sum_{n=1}^{N_m} p_{n,m} \le P_{\max}, && \text{(2a)} \\ & R_{n,m} \ge r_{n,m}, \forall n,m, && \text{(2b)} \\ & p_{n,m} \ge 0, \quad \forall n,m, && \text{(2c)} \end{aligned}$$

and the WEEM problem:

$$\begin{aligned} \mathcal{P}_1^{\rm WEEM}: \underset{\{\mathbf{p}_m\}_{m=1}^M}{\text{maximize}} \ & \frac{\sum\limits_{m=1}^{M}\sum\limits_{n=1}^{N_m} w_{n,m}R_{n,m}}{P_{\rm T}+\sum\limits_{m=1}^{M}\sum\limits_{n=1}^{N_m} p_{n,m}} \\ \text{subject to} \ & \sum_{m=1}^{M}\sum_{n=1}^{N_m} p_{n,m} \le P_{\max}, && \text{(3a)} \\ & R_{n,m} \ge r_{n,m}, \forall n,m, && \text{(3b)} \\ & p_{n,m} \ge 0, \quad \forall n,m, && \text{(3c)} \end{aligned}$$

where $w_{n,m}$ is the weight of $\text{UE}_{n,m}$, $\{\mathbf{p}_m\}_{m=1}^M$ is the set of power allocation vectors and $\mathbf{p}_m = [p_{1,m}, p_{2,m}, \ldots, p_{N_m,m}]^{\rm T}$ is the vector of allocated powers to the users in group m. Constraints (2a) and (3a) represent the transmitted power constraints, where $P_{\max}$ is the power budget of the BS, $P_{\rm T}$ represents the circuit power consumption and (2c), (3c) impose the nonnegativity constraints. Notably, the QoS constraints (2b) and (3b) with QoS threshold $r_{n,m} \ge 0$ for each n and m, are employed to ensure that the minimal data rate of each user is satisfied. This approach is unusual as it rarely coexists with user weights in existing works.

III. Optimal Power Allocation for Weighted Sum Rate Maximization

In this section, we first decompose the MG-NOMA WSRM problem into two layers, the upper layer for inter-group power budget allocation and the lower layer for intra-group power allocation. Then, we divide each group with several independent subgroups with independent power allocation strategy and derive an optimal closed-form intra-group power allocation scheme. Then, we transform the inter-group power budget allocation problem into a convex problem and seek the optimal power allocation solution of the MG-NOMA WSRM problem.

A. Reformulation of the WSRM Problem

To solve $\mathcal{P}_1^{\rm WSRM}$, we first simplify (1) and introduce the auxiliary variable $\mathbf{q}_m = [q_{1,m}, q_{2,m}, \ldots, q_{N_m,m}]^{\rm T}$, where $q_{n,m} = \sum_{i=n}^{N_m} p_{i,m}$. Thus conversely, we have

$$p_{n,m} = \begin{cases} q_{n,m} - q_{n+1,m}, & n < N_m, \\ q_{n,m}, & n = N_m. \end{cases} \tag{4}$$

Consequently, the formulation of rate $R_{n,m}$ in (1) can be transformed into:

$$\begin{aligned} R_{n,m} &= B_c \log\left(\frac{1+\alpha_{n,m}\sum\limits_{i=n}^{N_m} p_{i,m}}{1+\alpha_{n,m}\sum\limits_{i>n} p_{i,m}}\right) \\ &= \begin{cases} B_c \log\left(\frac{1+\alpha_{n,m}q_{n,m}}{1+\alpha_{n,m}q_{n+1,m}}\right), & n < N_m, \\ B_c \log\left(1+\alpha_{N_m,m}q_{N_m,m}\right), & n = N_m. \end{cases} \end{aligned}$$

In this way, the weighted sum rate is equivalent to:

$$\begin{aligned} \sum_{n=1}^{N_m} w_{n,m}R_{n,m} &= \sum_{n=1}^{N_m-1} w_{n,m}B_c \log\left(\frac{1+\alpha_{n,m}q_{n,m}}{1+\alpha_{n,m}q_{n+1,m}}\right) \\ &\quad + w_{N_m,m}B_c \log\left(1+\alpha_{N_m,m}q_{N_m,m}\right) \\ &= \sum_{n=1}^{N_m} f_{n,m}\left(q_{n,m}\right), \end{aligned}$$

where $f_{1,m}\left(q_{1,m}\right) \triangleq w_{1,m}B_c \log\left(1+\alpha_{1,m}q_{1,m}\right)$ and for $n > 1$, we have

$$\begin{aligned} f_{n,m}\left(q_{n,m}\right) &= w_{n,m}B_c \log\left(1+\alpha_{n,m}q_{n,m}\right) \\ &\quad - w_{n-1,m}B_c \log\left(1+\alpha_{n-1,m}q_{n,m}\right). \end{aligned} \tag{5}$$

Furthermore, the nonnegativity constraints require $q_{n,m} \ge q_{n+1,m}$ for $n < N_m$ and $q_{N_m,m} \ge 0$. The QoS constraint is equal to $\eta_{n,m}q_{n,m} - \beta_{n,m} \ge q_{n+1,m}$ for $n < N_m$ and $q_{N_m,m} \ge \theta_{N_m,m}$, where $\eta_{n,m} \triangleq 2^{-\frac{r_{n,m}}{B_c}}$, $\beta_{n,m} \triangleq \left(1-\eta_{n,m}\right)/\alpha_{n,m}$, $\theta_{n,m} \triangleq \left(2^{\frac{r_{n,m}}{B_c}}-1\right)/\alpha_{n,m}$. Regarding the nonnegativity constraints, it can be verified that $q_{N_m,m} \ge \theta_{N_m,m}$ implies $q_{N_m,m} \ge 0$ since $\theta_{N_m,m} \ge 0$. If $\eta_{n,m}q_{n,m} - \beta_{n,m} \ge q_{n+1,m}$ holds, we have $q_{n,m} - q_{n+1,m} \ge \left(1-\eta_{n,m}\right)q_{n,m} + \beta_{n,m} \ge 0$ for $n < N_m$. Thus, the nonnegativity constraints of $\mathcal{P}_1^{\rm WSRM}$ and $\mathcal{P}_1^{\rm WEEM}$ can be omitted if the QoS constraints are satisfied.

To tackle the power allocation in MG-NOMA, we introduce the power budgets $\{P_m\}_{m=1}^M$, where P_m represents the power budget for channel m. Suppose that the power budgets $\{P_m\}_{m=1}^M$ are given, then $\mathcal{P}_1^{\rm WSRM}$ can be equivalently decomposed into two layers, the upper layer for power budget allocation among groups and the lower layer for power allocation within each group. We first solve the SG-NOMA power allocation problem for group m, which is formulated as

$$\mathcal{P}_{2,m}^{\mathrm{WSRM}}: \underset{\mathbf{q}_m}{\text{maximize}} \quad \sum_{n=1}^{N_m} f_{n,m}(q_{n,m})$$
$$\text{subject to} \quad q_{1,m} = P_m, \tag{6a}$$
$$\eta_{n,m} q_{n,m} - \beta_{n,m} \geq q_{n+1,m}, n < N_m, \tag{6b}$$
$$q_{N_m,m} \geq \theta_{N_m,m}. \tag{6c}$$

The main difficulty in solving the WSRM problem in SG-NOMA lies in the coexistence of user weight and QoS constraint. User weight alters the characteristics of the weighted sum rate and individual optimums for different users may conflict with each other. Meanwhile, QoS constraints will make it more complicated, which is apparently difficult to find its solution. To address it, we introduce the following QoS transformation.

Lemma 1. *Suppose that $R_{i,m} = r_{i,m}$ holds for any $i = n, \ldots, l-1$, we have*

$$q_{i,m} = \xi_{n,i,m} q_{n,m} - \rho_{n,i,m}, \tag{7}$$

for $i = n, \ldots, l$, where $\xi_{n,i,m} \triangleq \prod_{j=n}^{i-1} \eta_{j,m}$, $\rho_{n,i,m} \triangleq \sum_{j=n}^{i-1} \xi_{j+1,i,m} \beta_{j,m}$ for $n < i$ and $\xi_{n,i,m} \triangleq 1$, $\rho_{n,i,m} \triangleq 0$ for $i = n$.

Proof. See Appendix A. □

In addition, the following result provides a necessary and sufficient condition on the feasibility of SG-NOMA WSRM.

Proposition 1. *$\mathcal{P}_{2,m}^{\mathrm{WSRM}}$ is feasible if and only if*

$$C1: \quad P_m \geq \Psi_{1,m}, \tag{8}$$

where $\Psi_{n,m} = (\theta_{N_m,m} + \rho_{n,N_m,m}) / \xi_{n,N_m,m}$.

Proof. See Appendix B. □

We define $\Psi_{1,m}$ as the minimum required power budget for group m. Now, we assume that the feasibility condition C1 in Proposition 1 holds and seek the optimal solution to $\mathcal{P}_{2,m}^{\mathrm{WSRM}}$.

B. Characteristics of the Weighted Sum Rate

In this subsection, we analyze the characteristics of the weighted sum rate. Clearly, a feasible solution for SG-NOMA is optimal if it achieves respective maximizer for each user. However, the characteristics of $f_{n,m}(q_{n,m})$ may not be that ideal, where the power allocation of different users are interdependent. To obtain the optimal power allocation in SG-NOMA, we consider dividing the group into several independent subgroups, which have independent power allocation strategy and their maximizers can coexist with each other. Therefore, we shall find the users with interdependent power allocation and group them together, which can be achieved by analyzing the weighted sum rate. Then, we introduce the auxiliary integrated subgroup $\mathcal{S}_{n,l,m}$ as defined below, where $\mathcal{S}_{n,l,m} = \{n, \ldots, l\}$ with $n \leq l$ is the collection of successive user indices from the n-th user to the l-th user in group m. The combination strategy depends on the properties of the function. Therefore, it is essential to analyze the function of the subgroup and we first provide a rigorous definition of the subgroup as follows.

Definition 1. A subgroup $\mathcal{S}_{n,l,m}$ is integrated if the QoS constraint (6b) always holds with equality at the optimal solution to $\mathcal{P}_{2,m}^{\mathrm{WSRM}}$, i.e., $R_{i,m} = r_{i,m}$, for $i = n, \ldots, l-1$.

Since QoS constraints for $\mathcal{S}_{n,l,m}$ take the equality for $i = n, \ldots, l-1$, we have $\eta_{i,m} q_{i,m} - \beta_{i,m} = q_{i+1,m}$ for $i = n, \ldots, l-1$. In this case, there always exists the affine relationship between adjacent variables $q_{i,m}$ and $q_{i+1,m}$ for $i = n, \ldots, l-1$. Through iteratively transformation, all $q_{i,m}$ for $i \in \mathcal{S}_{n,l,m}$ can be rewritten by a linear expression of $q_{n,m}$, i.e., $\xi_{n,i,m} q_{n,m} - \rho_{n,i,m} = q_{l+1,m}$. All users belong to $\mathcal{S}_{n,l,m}$ can also be viewed as a virtual user with variable $q_{n,m}$. Furthermore, the $q_{i,m}$ in the corresponding function $f_{i,m}(q_{i,m})$ can be replaced by an affine result of $q_{n,m}$, i.e., $f_{i,m}(\xi_{n,i,m} q_{n,m} - \rho_{n,i,m})$. Next, the weighted sum rate of $\mathcal{S}_{n,l,m}$ can be equivalently transformed into a function of $q_{n,m}$, which is defined as $g_{n,l,m}(q_{n,m}) = \sum_{i=n}^{l} f_{i,m}(\xi_{n,i,m} q_{n,m} - \rho_{n,i,m})$. The expression $g_{n,l,m}(q_{n,m})$ is detailed in (9) and (10) on the top of next page for $n = 1$ and $n > 1$, respectively. Note that a single user can also be viewed as an integrated subgroup.

As mentioned above, the integrated subgroup $\mathcal{S}_{n,l,m}$ may be combined with adjacent subgroups according to the characteristics of $g_{n,l,m}(q_{n,m})$. The combining criteria can be determined by the property of $g_{n,l,m}(q_{n,m})$. For example, suppose that there exist two adjacent subgroups $\mathcal{S}_{n,l,m}$, $\mathcal{S}_{l+1,k,m}$ are both integrated and $g_{l+1,k,m}(q_{l+1,m})$ is increasing with respect to $q_{l+1,m}$. In this case, the $q_{l+1,m}$ always take the maximum value and thus $\eta_{l,m} q_{l,m} - \beta_{l,m} \geq q_{l+1,m}$ always holds with equality at the optimal solution to $\mathcal{P}_{2,m}^{\mathrm{WSRM}}$. In this case, $\mathcal{S}_{n,l,m}$ and $\mathcal{S}_{l+1,k,m}$ shall be combined and form a newly integrated subgroup $\mathcal{S}_{n,k,m}$ since the other requirements of $\eta_{i,m} q_{i,m} - \beta_{i,m} = q_{i+1,m}$ for $\mathcal{S}_{n,k,m}$ are met with the integrated characteristics of $\mathcal{S}_{n,l,m}$ and $\mathcal{S}_{l+1,k,m}$. After iterative combining until termination, the group will be divided into several independent subgroups since any interdependent subgroups will be combined together. The subgroup combining method and power allocation strategy are provided in the following results.

Denote $\Omega_{n,l,m}$ as the maximizer of $g_{n,l,m}(q_{n,m})$, assuming the maximizer exists. We then proceed to analyze $g_{n,l,m}(q_{n,m})$ and, based on this, develop the subgroup combining strategy. To investigate the characteristics of $g_{n,l,m}(q_{n,m})$, we shall consider the relation between $\xi_{n,l,m}$ and $\rho_{n,l,m}$. Given $\alpha_{1,m} \leq \cdots \leq \alpha_{N_m,m}$, it can be verified that $\xi_{n,l,m} - 1 + \alpha_{l,m} \rho_{n,l,m} \geq 0$ holds, which is obtained by

$$\begin{aligned}
&\xi_{n,l,m} - 1 + \alpha_{l,m} \rho_{n,l,m} \\
&= \xi_{n,l,m} - 1 + \alpha_{l,m} \sum_{i=n}^{l-1} \xi_{i+1,l,m} \frac{1 - \eta_{i,m}}{\alpha_{i,m}} \\
&\geq \xi_{n,l,m} - 1 + \sum_{i=n}^{l-1} \xi_{i+1,l,m} (1 - \eta_{i,m}) \\
&= \xi_{n,l,m} - \xi_{n,l,m} + \xi_{n+1,l,m} - \ldots - \xi_{l-1,l,m} + \xi_{l,l,m} - 1 \\
&= \xi_{l,l,m} - 1 = 0.
\end{aligned} \tag{12}$$

$$\begin{aligned} g_{1,l,m}\left(q_{n,m}\right) =& \sum_{i=1}^{l} f_{i,m}\left(q_{i,m}\right) = w_{l,m}B_c \log\left(1+\alpha_{l,m}q_{l,m}\right) + \sum_{i=1}^{l-1} w_{i,m}r_{i,m} \\ =& w_{l,m}B_c \log\left(1+\alpha_{l,m}\xi_{1,l,m}q_{1,m} - \alpha_{l,m}\rho_{1,l,m}\right) + \sum_{i=1}^{l-1} w_{i,m}r_{i,m}. \end{aligned} \tag{9}$$

$$\begin{aligned} g_{n,l,m}\left(q_{n,m}\right) =& \sum_{i=n}^{l} f_{i,m}\left(q_{i,m}\right) \\ =& w_{l,m}B_c \log\left(1+\alpha_{l,m}q_{l,m}\right) - w_{n-1,m}B_c \log\left(1+\alpha_{n-1,m}q_{n,m}\right) + \sum_{i=n}^{l-1} w_{i,m}r_{i,m} \\ =& w_{l,m}B_c \log\left(1+\alpha_{l,m}\xi_{n,l,m}q_{n,m} - \alpha_{l,m}\rho_{n,l,m}\right) - w_{n-1,m}B_c \log\left(1+\alpha_{n-1,m}q_{n,m}\right) + \sum_{i=n}^{l-1} w_{i,m}r_{i,m}. \end{aligned} \tag{10}$$

$$\Gamma_{n,l,m} = \frac{w_{l,m}\alpha_{l,m}\xi_{n,l,m} - w_{n-1,m}\alpha_{n-1,m} + w_{n-1,m}\alpha_{n-1,m}\alpha_{l,m}\rho_{n,l,m}}{\left(w_{n-1,m} - w_{l,m}\right)\alpha_{n-1,m}\alpha_{l,m}\xi_{n,l,m}}. \tag{11}$$

The combining target and criteria can determined by the property of the function. Thus, we provide the monotonicity conditions of $g_{n,l,m}\left(q_{n,m}\right)$ in the following results.

Lemma 2. *Given $\alpha_{1,m} \leq \cdots \leq \alpha_{N_m,m}$, $g_{n,l,m}\left(q_{n,m}\right)$, is monotonically nondecreasing in $q_{n,m}$ if the following condition is met for $n > 1$:*

$$C2: \quad w_{l,m} \geq w_{n-1,m}. \tag{13}$$

Then, $\eta_{n-1,m}q_{n-1,m} - \beta_{n-1,m} = q_{n,m}$ is met at the optimal solution of $\mathcal{P}_{2,m}^{\mathrm{WSRM}}$.

Proof. See Appendix C. □

In this case, we define $\mathcal{S}_{n,l,m}$ as a nondecreasing subgroup. This leads to the necessity of combining $\mathcal{S}_{n,l,m}$ with the preceding subgroup if $n > 1$, specifically, the subgroup encompassing $\mathrm{UE}_{n-1,m}$ since $\eta_{n-1,m}q_{n-1,m} - \beta_{n-1,m} = q_{n,m}$. As for $n = 1$, one can observe that $g_{1,l,m}\left(q_{1,m}\right)$ is monotonically nondecreasing in $q_{1,m}$ according to (9). Thus, the integrated subgroup $\mathcal{S}_{1,l,m}$ is nondecreasing. Next, we analyze the monotonicity of $g_{n,l,m}\left(q_{n,m}\right)$ if C2 is violated, which is given below.

Lemma 3. *Given $\alpha_{1,m} \leq \cdots \leq \alpha_{N_m,m}$, there exists a maximizer $\Omega_{n,l,m} = \max\left\{\Gamma_{n,l,m}, \Psi_{n,m}\right\}$ for $g_{n,l,m}\left(q_{n,m}\right)$ if the following condition holds for $n > 1$:*

$$C3: \quad w_{l,m} < w_{n-1,m}, \tag{14}$$

where $\Gamma_{n,l,m}$ is given by (11).

Proof. See Appendix D. □

In this case, we define $\mathcal{S}_{n,l,m}$ as a subgroup which has a maximizer $\Omega_{n,l,m}$. Hence, for any integrated subgroup $\mathcal{S}_{n,l,m}$, either C2 or C3 will be met and it is nondecreasing or has a maximizer $\Omega_{n,l,m}$.

C. Optimal Power Allocation for Single-Group WSRM Problem

In this subsection, we formulate the closed-form optimal power allocation for the single-group WSRM problem, i.e., $\mathcal{P}_{2,m}^{\mathrm{WSRM}}$. As summarized in Section III-B, the weighted sum rate function for any integrated subgroup is either nondecreasing or has a maximizer. According to Lemma 2, two adjacent integrated subgroups can be combined if the latter is nondecreasing. Furthermore, when two adjacent subgroups $\mathcal{S}_{n,l,m}$ and $\mathcal{S}_{l+1,k,m}$ reach their respective maximizers, i.e., $q_{n,m} = \Omega_{n,l,m}$ and $q_{l+1,m} = \Omega_{l+1,k,m}$, the solution will be infeasible if the QoS constraint $\eta_{l,m}q_{l,m} - \beta_{l,m} \geq q_{l+1,m}$ is violated, and the criteria for potential combining is provided as below.

Lemma 4. *Suppose that two adjacent integrated subgroups $\mathcal{S}_{n,l,m}$ and $\mathcal{S}_{l+1,k,m}$ which have respective maximizers, these two subgroups shall be combined to yield a newly integrated subgroup $\mathcal{S}_{n,k,m}$ unless the following condition holds:*

$$C4: \quad \xi_{n,l+1,m}\Omega_{n,l,m} - \rho_{n,l+1,m} > \Omega_{l+1,k,m}. \tag{15}$$

Proof. See Appendix E. □

After iteratively combining until termination, we define $\{\mathcal{U}_{t,m}\}_{t=1}^{T_m}$ as an independent group dividing scheme and T_m as the number of the subgroups in $\{\mathcal{U}_{t,m}\}_{t=1}^{T_m}$. Only the initial subgroup in $\{\mathcal{U}_{t,m}\}_{t=1}^{T_m}$ remains nondecreasing, as any other nondecreasing subgroup would be combined with its preceding subgroup according to Lemma 2. The remaining subgroups have respective maximizers, and C4 holds between adjacent subgroups. Otherwise, the adjacent subgroups will be combined according to Lemma 4.

Since the affine relationship between the users in the same integrated subgroup has been confirmed above, the power allocation of users can be excluded from consideration except the initial user in each subgroup. For convenience of notation, we use $J_{t,m}$ and $L_{t,m}$ to represent the initial and final user indices in $\mathcal{U}_{t,m}$, respectively. Thus, we have $\mathcal{U}_{t,m} = \{J_{t,m}, J_{t,m}+1, \ldots, L_{t,m}\}$, $J_{1,m} = 1$, $L_{T_m,m} = N_m$. One can observe that $L_{t-1,m}+1 = J_{t,m}$ holds for $t > 1$ due to the adjacency between $\mathcal{U}_{t-1,m}$ and $\mathcal{U}_{t,m}$. Furthermore, C4 can be equivalently reinterpreted as $\psi_{t,m}\Phi_{t,m} - \delta_{t,m} > \Phi_{t+1,m}$ for $t = 2, \ldots, T_m - 1$, where $\Phi_{t,m} = \Omega_{J_{t,m},L_{t,m},m}$ for $t > 1$, and $\psi_{t,m} \triangleq \xi_{J_{t,m},J_{t+1,m},m}$, $\delta_{t,m} \triangleq \rho_{J_{t,m},J_{t+1,m},m}$

$$\begin{aligned}\psi_{t,m}\Phi_{t,m}-\delta_{t,m} &= \xi_{J_{t,m},J_{t+1,m},m}\left(\xi_{1,J_{t,m},m}\Lambda_{t,m}-\rho_{1,J_{t,m},m}\right)-\rho_{J_{t,m},J_{t+1,m},m}\\ &= \xi_{1,J_{t+1,m},m}\Lambda_{t,m}-\xi_{J_{t,m},J_{t+1,m},m}\sum_{j=1}^{J_{t,m}-1}\xi_{j+1,J_{t,m},m}\beta_{j,m}-\sum_{j=J_{t,m}}^{J_{t+1,m}-1}\xi_{j+1,J_{t+1,m},m}\beta_{j,m}\\ &= \xi_{1,J_{t+1,m},m}\Lambda_{t,m}-\sum_{j=1}^{J_{t+1,m}-1}\xi_{j+1,J_{t,m},m}\beta_{j,m}\\ &= \xi_{1,J_{t+1,m},m}\Lambda_{t,m}-\rho_{1,J_{t+1,m},m}\\ &> \Phi_{t+1,m}=\xi_{1,J_{t+1,m},m}\Lambda_{t+1,m}-\rho_{1,J_{t+1,m},m}. \end{aligned} \tag{16}$$

Algorithm 1 Group Dividing Method for Group m

1: **Initialization:** $\{\mathcal{S}_t\}_{t=1}^{T_m}$, where $\mathcal{S}_t=\{t\}$ and $T_m=N_m$.
2: **while** $\{\mathcal{S}_t\}_{t=1}^{T_m}$ is not independent **do**
3:   **for** $t=2$ **to** T_m **do**
4:     **if** $\mathcal{S}_t$ is nondecreasing **then**
5:       $\mathcal{S}_t$ is combined with $\mathcal{S}_{t-1}$: $\mathcal{S}_{t-1}=\mathcal{S}_{t-1}\cup\mathcal{S}_t$; $\mathcal{S}_i=\mathcal{S}_{i+1}$ for $i\geq t$; $T_m=T_m-1$; break;
6:     **else if** C4 does not holds between $\mathcal{S}_t$ and $\mathcal{S}_{t-1}$ **then**
7:       $\mathcal{S}_t$ is combined with $\mathcal{S}_{t-1}$: $\mathcal{S}_{t-1}=\mathcal{S}_{t-1}\cup\mathcal{S}_t$; $\mathcal{S}_i=\mathcal{S}_{i+1}$ for $i\geq t$; $T_m=T_m-1$; break;
8:     **end if**
9:   **end for**
10: **end while**
11: Set $\mathcal{U}_{t,m}=\mathcal{S}_t$ for $t=1,\ldots,T_m$;
12: **Output:** $\{\mathcal{U}_{t,m}\}_{t=1}^{T_m}$.

---

for $t=2,\ldots,T_m-1$. In this case, the users in group m can be divided into T_m subgroups with independent optimal power strategy, otherwise there will be two adjacent subgroups to be combined together according to Lemmas 2, 3 and 4. The process of group dividing for group m is described in Algorithm 1. Now, we consider the independent power allocation strategies for the subgroups in $\{\mathcal{U}_{t,m}\}_{t=1}^{T_m}$, and the optimal solution can be characterized in closed-form.

Theorem 1. *The optimal solution to $\mathcal{P}_{2,m}^{\mathrm{WSRM}}$ is given by $q_{n,m}^{\star}=\xi_{1,n,m}P_m-\rho_{1,n,m}$ for $n\in\mathcal{U}_{1,m}$ and*

$$q_{n,m}^{\star}=\min\left\{\xi_{1,n,m}P_m-\rho_{1,n,m},\xi_{J_{t,m},n,m}\Phi_{t,m}-\rho_{J_{t,m},n,m}\right\},$$

for $n\in\mathcal{U}_{t,m}$, $t>1$, where $\{\mathcal{U}_{t,m}\}_{t=1}^{T_m}$ is the independent group dividing scheme for group m.

Proof. See Appendix F. □

Before proceeding to the main results, we start from the following toy example to gain some insights.

Example 1. We investigate a single group m, with $N_m=3$ users within the group m. For convenience of discussion, the group index in Example 1 is omitted. Suppose that the channel gain set $\mathcal{C}=\{1,4,16\}$, QoS constraint set $\mathcal{V}=\{1,1,1\}$ and weight set $\mathcal{W}=\{1,2,3\}$, where the n-th element in $\mathcal{C}$, $\mathcal{V}$ and $\mathcal{W}$ represents the channel gain α_n, QoS constraint r_n and user weight w_n for UE_n, respectively. In the initialization phase at line 1 of Algorithm 1, each user can be viewed as a subgroup, i.e., $T_m=3$, $\mathcal{S}_1^{(0)}=\{1\}$, $\mathcal{S}_2^{(0)}=\{2\}$, $\mathcal{S}_3^{(0)}=\{3\}$. It follows from Lemma 2 that $\mathcal{S}_2^{(0)}$ is nondecreasing since $w_1\leq w_2$. In this case, the condition at line 4 of Algorithm 1 is satisfied and the line 5 of Algorithm 1 will be implemented. Then, $\mathcal{S}_2^{(0)}$ will be combined with $\mathcal{S}_1^{(0)}$, i.e., $\mathcal{S}_1^{(1)}=\mathcal{S}_1^{(0)}\cup\mathcal{S}_2^{(0)}=\{1,2\}$, the original $\mathcal{S}_3^{(0)}$ is adjusted to the newly $\mathcal{S}_2^{(1)}=\mathcal{S}_3^{(0)}=\{3\}$ and $T_m=T_m-1=2$. In the next loop, according to Lemma 2, $\mathcal{S}_2^{(1)}$ is also nondecreasing since $w_2\leq w_3$. Similarly, $\mathcal{S}_2^{(1)}$ will be combined with newly $\mathcal{S}_1^{(1)}$, i.e., $\mathcal{S}_1^{(2)}=\mathcal{S}_1^{(1)}\cup\mathcal{S}_2^{(1)}=\{1,2,3\}$ and $T_m=T_m-1=1$. Finally, the three user in group m form a subgroup and we have $T_m=1$, $\mathcal{U}_{1,m}=\{1,2,3\}$. Then, the closed-form optimal power allocation is given by $q_{n,m}^{\star}=\xi_{1,n,m}P_m-\rho_{1,n,m}$ for $n=1,2,3$.

Therefore, the users in $\mathcal{P}_{2,m}^{\mathrm{WSRM}}$ can be transformed into several integrated subgroups and the monotonicity of the corresponding weighted sum rate function is characterized by Lemma 2 and Lemma 3. Furthermore, the independent group dividing scheme is obtained via Algorithm 1 and the optimal solution to $\mathcal{P}_{2,m}^{\mathrm{WSRM}}$ is provided through Theorem 1.

D. Optimal Power Allocation for Multi-Group WSRM Problem

In Section III-C, we have successfully characterized the optimal weighted sum rate analytically and formulated the optimal intra-group power allocation in group m for $\mathcal{P}_{2,m}^{\mathrm{WSRM}}$. Drawing from the insights derived from Lemmas 2, 3 and 4, we can segregate users in group m into T_m integrated subgroups. The optimal power allocation of $\mathcal{P}_{2,m}^{\mathrm{WSRM}}$ is then elucidated in Theorem 1. Regarding the subgroup $\mathcal{U}_{t,m}$ for $t>1$, the optimal solution achieves

$$q_{J_{t,m},m}^{\star}=\begin{cases}\xi_{1,J_{t,m},m}P_m-\rho_{1,J_{t,m},m}, & P_m\geq\Lambda_{t,m},\\ \Phi_{t,m}, & P_m<\Lambda_{t,m},\end{cases}$$

where $\Lambda_{t,m}\triangleq\left(\Phi_{t,m}+\rho_{1,J_{t,m},m}\right)/\xi_{1,J_{t,m},m}$ for $t>1$. Conversely, $\Phi_{t,m}=\xi_{1,J_{t,m},m}\Lambda_{t,m}-\rho_{1,J_{t,m},m}$ for $t>1$. For convenience of notation, we denote the lower bound of P_m as $\Lambda_{T_m+1,m}=\Psi_{1,m}$. It follows from (16) that $\Lambda_{t,m}>\Lambda_{t+1,m}$ is satisfied for $t=2,\ldots,T_m-1$. In addition, we have $\Lambda_{T_m,m}\geq\Lambda_{T_m+1,m}$ since $\Phi_{t,m}=\Omega_{J_{t,m},L_{t,m},m}\geq\Psi_{J_{t,m},m}$ according to Lemma 3.

In line with the variable transformation detailed in Lemma 1, the weighted sum rate function of the integrated subgroup

$$\begin{aligned}\frac{\hat{w}_{t,m}\hat{\alpha}_{t,m}\hat{\xi}_{t,m}B_c}{1+\hat{\alpha}_{t,m}\hat{\xi}_{t,m}\Lambda_{t,m}-\hat{\alpha}_{t,m}\hat{\rho}_{t,m}} &= \frac{w_{L_{t,m},m}\alpha_{L_{t,m},m}\xi_{1,L_{t,m},m}B_c}{1+\alpha_{L_{t,m},m}\xi_{1,L_{t,m},m}\Lambda_{t,m}-\alpha_{L_{t,m},m}\rho_{1,L_{t,m},m}} \\ &= \frac{w_{L_{t,m},m}\alpha_{L_{t,m},m}\xi_{1,L_{t,m},m}B_c}{1+\alpha_{L_{t,m},m}\xi_{J_{t,m},L_{t,m},m}\Phi_{t,m}-\alpha_{L_{t,m},m}\rho_{J_{t,m},L_{t,m},m}} = \frac{w_{L_{t-1,m},m}\alpha_{L_{t-1,m},m}\xi_{1,J_{t,m},m}B_c}{1+\alpha_{J_{t-1,m},m}\Phi_{t,m}} \\ &= \frac{w_{L_{t-1,m},m}\alpha_{L_{t-1,m},m}\xi_{1,J_{t,m},m}B_c/\eta_{L_{t-1,m},m}}{1+\alpha_{L_{t-1,m},m}\xi_{L_{t-1,m},m}\Lambda_{t,m}-\alpha_{L_{t-1,m},m}\rho_{L_{t-1,m},m}} = \frac{\hat{w}_{t-1,m}\hat{\alpha}_{t-1,m}\hat{\xi}_{t-1,m}B_c}{1+\hat{\alpha}_{t-1,m}\hat{\xi}_{t-1,m}\Lambda_{t,m}-\hat{\alpha}_{t-1,m}\hat{\rho}_{t-1,m}}.\end{aligned} \tag{17}$$

$\mathcal{U}_{t,m}$ can be equivalently transformed into a function of $q_{J_{t,m},m}$, which is given by

$$\begin{aligned}&\phi_{t,m}\left(q_{J_{t,m},m}\right) \\ &= \sum_{n=J_{t,m}}^{L_{t,m}} f_{n,m}\left(q_{n,m}\right) \\ &= \sum_{n=J_{t,m}}^{L_{t,m}} f_{n,m}\left(\xi_{J_{t,m},n,m}q_{J_{t,m},m}-\rho_{J_{t,m},n,m}\right).\end{aligned} \tag{18}$$

If $P_m < \Lambda_{t,m}$ for $t>1$, the optimal solution achieves $R^\star_{n,m} = r_{n,m}$ for $n < L_{t,m}$ and we have

$$\sum_{n=1}^{L_{t,m}} f_{n,m}\left(q_{n,m}\right) = \sum_{n<L_{t,m}} w_{n,m}r_{n,m} + G_{t,m}\left(P_m\right),$$

where

$$G_{t,m}\left(P_m\right) \triangleq \hat{w}_{t,m}B_c \log\left(1+\hat{\alpha}_{t,m}\hat{\xi}_{t,m}P_m-\hat{\alpha}_{t,m}\hat{\rho}_{t,m}\right),$$

$\hat{w}_{t,m} \triangleq w_{L_{t,m},m}$, $\hat{\alpha}_{t,m} \triangleq \alpha_{L_{t,m},m}$, $\hat{\xi}_{t,m} \triangleq \xi_{1,L_{t,m},m}$ and $\hat{\rho}_{t,m} \triangleq \rho_{1,L_{t,m},m}$.

As for $P_m \geq \Lambda_{t,m}$, we have $q^\star_{J_{i,m},m} = \Phi_{i,m}$ and $\sum_{n=J_{i,m}}^{L_{i,m}} f_{n,m}\left(q_{n,m}\right) = \phi_{i,m}\left(\Phi_{i,m}\right)$ for $i \geq t$. For convenience of comparison between P_m and $\Lambda_{t,m}$, we introduce auxiliary index $t^\star = \arg\min\{t \mid P_m \geq \Lambda_{t,m}\} - 1$. Since $\Lambda_{t,m} > \Lambda_{t+1,m}$ holds for $t = 2,\ldots,T_m-1$, we have $P_m \geq \Lambda_{i,m}$ if $i \geq t^\star$ and $P_m < \Lambda_{i,m}$ if $i < t^\star$. In this way, the weighted sum rate function for group m is given by

$$\begin{aligned}F_m\left(P_m\right) &= \sum_{n=1}^{N_m} f_{n,m}\left(q^\star_{n,m}\right) \\ &= \phi_{1,m}\left(P_m\right) + \sum_{t=2}^{T_m}\phi_{t,m}\left(q^\star_{J_{t,m},m}\right) \\ &= G_{t^\star,m}\left(P_m\right) + \Upsilon_{t^\star,m},\end{aligned} \tag{19}$$

where $\Upsilon_{t,m} = \sum_{n<L_{t,m}} w_{n,m}r_{n,m} + \sum_{i>t}\phi_{i,m}\left(\Phi_{i,m}\right)$ is a constant.

Obviously, $F_m\left(P_m\right)$ is a piecewise and continuous function. In addition, $F_m\left(P_m\right)$ is differentiable within each interval. As for any junction bridging adjacent intervals $\Lambda_{t,m}$, we have $\Lambda_{t,m} > \Psi_{1,m}$ and thus $\Omega_{J_{t,m},L_{t,m},m} = \Gamma_{J_{t,m},L_{t,m},m}$. It can be verified that $F_m\left(P_m\right)$ is differentiable at the junction $\Lambda_{t,m}$ according to (17). For convenience, the constants $\ln 2$ in the denominators is omitted.

Consequently, $F_m\left(P_m\right)$ is differentiable and its derivative can be formulated as

$$\frac{\mathrm{d}F_m}{\mathrm{d}P_m} = \frac{\hat{w}_{t^\star,m}\hat{\alpha}_{t^\star,m}\hat{\xi}_{t^\star,m}B_c}{\left(1+\hat{\alpha}_{t^\star,m}\hat{\xi}_{t^\star,m}P_m-\hat{\alpha}_{t^\star,m}\hat{\rho}_{t^\star,m}\right)\ln 2}. \tag{20}$$

It is evident that the derivative of $F_m\left(P_m\right)$ is continuous and monotonically decreasing in P_m, which implies that $F_m\left(P_m\right)$ is a concave function (See Theorem 12.18 in [33]). As previously elaborated, the optimal solution of $\mathcal{P}_{2,m}^{\mathrm{WSRM}}$ can be achieved with given power budget P_m, and the optimal weighted sum rate for group m is characterized by $F_m\left(P_m\right)$ in (19). This characterization allows the optimal power allocation for maximizing weighted sum rate to be equivalently transformed into power budget allocation problem, which is delineated as

$$\begin{aligned}\mathcal{P}_3^{\mathrm{WSRM}}:\ \underset{\{P_m\}_{m=1}^M}{\text{maximize}}\ &\sum_{m=1}^M F_m\left(P_m\right) \\ \text{subject to}\ &\sum_{m=1}^M P_m \leq P_{\max}, &\text{(21a)} \\ &P_m \geq \Psi_{1,m}, \forall m. &\text{(21b)}\end{aligned}$$

Given the concavity of $F_m\left(P_m\right)$ and the linearity of constraints, $\mathcal{P}_3^{\mathrm{WSRM}}$ is proved to be a convex problem. As a result, it can be solved via standard convex optimization toolboxes, e.g., CVX [34]. Despite this, we are able to analytically characterize the optimal solution via exploiting the Karush-Kuhn-Tucker (KKT) conditions, as detailed in the following result.

Theorem 2. *The optimal solution to $\mathcal{P}_3^{\mathrm{WSRM}}$ is given by*

$$P_m = \max\left\{\frac{\tilde{w}_m(\lambda)B_c}{\lambda\ln 2} + \frac{\tilde{\rho}_m(\lambda)}{\tilde{\xi}_m(\lambda)} - \frac{1}{\tilde{\alpha}_m(\lambda)\tilde{\xi}_m(\lambda)}, \Psi_{1,m}\right\}, \tag{22}$$

where $\tilde{\alpha}_m(\lambda)$, $\tilde{w}_m(\lambda)$, $\tilde{\xi}_m(\lambda)$, $\tilde{\rho}_m(\lambda)$ are the related coefficients at different intervals, which are listed as follows:

1) $\tilde{\alpha}_m(\lambda) = \hat{\alpha}_{1,m}$, $\tilde{w}_m(\lambda) = \hat{w}_{1,m}$, $\tilde{\xi}_m(\lambda) = \hat{\xi}_{1,m}$, $\tilde{\rho}_m(\lambda) = \hat{\rho}_{1,m}$, if $\lambda \leq \frac{\mathrm{d}F_m}{\mathrm{d}\Lambda_{2,m}}$;

2) $\tilde{\alpha}_m(\lambda) = \hat{\alpha}_{t,m}$, $\tilde{w}_m(\lambda) = \hat{w}_{t,m}$, $\tilde{\xi}_m(\lambda) = \hat{\xi}_{t,m}$, $\tilde{\rho}_m(\lambda) = \hat{\rho}_{t,m}$, if $\frac{\mathrm{d}F_m}{\mathrm{d}\Lambda_{t,m}} < \lambda \leq \frac{\mathrm{d}F_m}{\mathrm{d}\Lambda_{t+1,m}}$ for $t>1$.

In the above, λ is chosen such that $\sum_{m=1}^M P_m = P_{\max}$.

Proof. See Appendix G. □

The fact that P_m is monotonically decreasing in λ suggests that the optimal λ satisfying $\sum_{m=1}^M P_m = P_{\max}$ can be efficiently found through a bisection method. Hence, our proposed OPA WSRM method in MG-NOMA systems is provided by Algorithm 2. After decomposing $\mathcal{P}_1^{\mathrm{WSRM}}$ into several problems, we analyze the characteristics of the weighted sum rate function for the subgroup in $\mathcal{P}_{2,m}^{\mathrm{WSRM}}$ and design a group

Algorithm 2 The Optimal Power Allocation for Weighted Sum Rate Maximization (OPA WSRM)

Obtain $\{\mathcal{U}_{t,m}\}_{t=1}^{T_m}$ using Algorithm 1, $\forall m$;
Set precision $\varepsilon_\lambda > 0$, $\lambda_u = \max\left\{\frac{\mathrm{d}F_1}{\mathrm{d}\Psi_{1,1}}, \ldots, \frac{\mathrm{d}F_M}{\mathrm{d}\Psi_{1,M}}\right\}$, $\lambda_d = 0$;
while $\lambda_u - \lambda_d > \varepsilon_\lambda$ **do**
  $\lambda = (\lambda_u + \lambda_d)/2$;
  Obtain P_m according to (22), $\forall m$.
  if $\sum_{m=1}^{M} P_m > P_{\max}$ **then**
    $\lambda_d = \lambda$;
  else
    $\lambda_u = \lambda$;
  end if
end while
Obtain $\{q_{n,m}^{\star}\}_{n=1}^{N_m}$ using Theorem 1, $\forall m$;
Obtain $\{p_{n,m}^{\star}\}_{n=1}^{N_m}$ according to (4), $\forall m$.

dividing method. Based on the group dividing scheme, we provide a closed-form solution and the weighted sum rate function for $\mathcal{P}_{2,m}^{\mathrm{WSRM}}$. Consequently, $\mathcal{P}_{1}^{\mathrm{WSRM}}$ can be transformed into a convex problem $\mathcal{P}_{3}^{\mathrm{WSRM}}$. By exploiting the KKT conditions of $\mathcal{P}_{3}^{\mathrm{WSRM}}$, the optimal solution to $\mathcal{P}_{3}^{\mathrm{WSRM}}$ is characterized in semi-closed form.

E. Complexity Analysis

In this subsection, we examine the complexity of different power allocation schemes. For each group m, the combining in Algorithm 1 is performed at most $N_m - 1$ iterations. Thus, the worst case complexity of Algorithm 1 for group m is $O(N_m)$. Since Algorithm 1 is implemented independently among M groups, the overall complexity in multi-group NOMA system is $\sum_{m=1}^{M} O(N_m) = O(N)$. Then, we discuss the computational complexity of our proposed OPA WSRM method in Algorithm 2. In each iteration of Algorithm 2, we obtain P_m for each group with given λ via (22), which is performed among M groups and requires $O(M)$ operations (additions and multiplications). Moreover, the complexities of intra-group and inter-group power allocation in Algorithm 2 are $O(N)$ and $O(I_{\mathrm{WSR}}M)$, respectively, where I_{WSR} is the number of iterations. Taking into account the complexity of Algorithm 1, the overall complexity of optimal power allocation for WSRM problem is $O(N + I_{\mathrm{WSR}}M)$.

We also provide the complexity analysis for several MG-NOMA sum rate maximization methods within the existing works. As detailed in [29], the Bisection method is performed with complexity $O(N + I_{\mathrm{B}}M)$, where I_{B} is the number of bisection iterations. In [28], the multi-carrier power control (MCPC) algorithm employs the dynamic programming method to maximize the weighted sum rate, where the power budget is allocated into J parts and the overall complexity is $O\left(N^3 + JN^2 + J^2M\right)$. In addition, a remaining power sharing (RPS) method has been developed in [21] and the overall complexity is $O(N)$.

IV. Optimal Power Allocation for Weighted Energy Efficiency Maximization

In this section, we seek the optimal solution of the MG-NOMA WEEM problem. We first transform the original WEEM problem from a nonconvex fractional programming problem into several convex subproblems. Then, by exploiting the KKT conditions, we obtain the optimal power allocation scheme for the WEEM problem in semi-closed form. Finally, we further show that an optimal solution can be derived in closed-form.

A. Optimal Power Allocation for the Multi-Group WEEM Problem

In this subsection, we present the optimal power allocation scheme for $\mathcal{P}_{1}^{\mathrm{WEEM}}$. Notably, given the total transmit power $\sum_{m=1}^{M} P_m$, $\mathcal{P}_{1}^{\mathrm{WEEM}}$ is equivalent to $\mathcal{P}_{1}^{\mathrm{WSRM}}$ in MG-NOMA systems. However, it is hard to determine the optimal total transmit power since the monotonicity of WEE with respect to the total transmit power in $\mathcal{P}_{1}^{\mathrm{WEEM}}$ is not established. Therefore, we propose a novel design to solve the WEEM problem. Based on the weighted sum rate function $F_m(P_m)$ derive in $\mathcal{P}_{2,m}^{\mathrm{WSRM}}$, we also transform the MG-NOMA WEEM problem into inter-group power budget allocation problem. By transforming the fractional programming problem into several subtractive form subproblems, we derive the optimal solution of the subproblems. On this basis, we are able to incorporate existing complex two-layer iterative solution into a low-complexity bisection based single-layer iteration method.

In this way, we first transform problem $\mathcal{P}_{1}^{\mathrm{WEEM}}$ into a power budget allocation problem, which is formulated as follows:

$$\mathcal{P}_{2}^{\mathrm{WEEM}}: \underset{\{P_m\}_{m=1}^{M}}{\text{maximize}} \quad \frac{\sum\limits_{m=1}^{M} F_m(P_m)}{P_{\mathrm{T}} + \sum\limits_{m=1}^{M} P_m}$$

$$\text{subject to} \quad \sum_{m=1}^{M} P_m \le P_{\max}, \tag{23a}$$

$$P_m \ge \Psi_{1,m}, \forall m. \tag{23b}$$

Clearly, $\mathcal{P}_{2}^{\mathrm{WEEM}}$ is a nonconvex fractional programming problem, and the challenge in solving it lies in the fraction form of the objective function. As proved in Section III-D, $F_m(P_m)$ is a concave function and the denominator of the objective function in $\mathcal{P}_{2}^{\mathrm{WEEM}}$ is an affine function. Hence, the optimal solution of $\mathcal{P}_{2}^{\mathrm{WEEM}}$ can be obtained via the well-known Dinkelbach's method [35], which has been used in many existing works of fractional programming problem. In Dinkelbach's algorithm, we transform $\mathcal{P}_{2}^{\mathrm{WEEM}}$ into an equivalent subtractive form as follows:

$$\mathcal{P}_{3}^{\mathrm{WEEM}}: \underset{\{P_m\}_{m=1}^{M}}{\text{maximize}} \quad G(\nu)$$

$$\text{subject to} \quad \sum_{m=1}^{M} P_m \le P_{\max}, \tag{24a}$$

$$P_m \ge \Psi_{1,m}, \forall m, \tag{24b}$$

This article has been accepted for publication in IEEE Transactions on Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2025.3616144

where $G(\nu) = \sum_{m=1}^{M} F_m(P_m) - \nu\left(P_{\mathrm{T}} + \sum_{m=1}^{M} P_m\right)$ and ν is a non-negative parameter. $\mathcal{P}_3^{\mathrm{WEEM}}$ can be viewed as an amended version of $\mathcal{P}_3^{\mathrm{WSRM}}$ with an extra term $-\nu\left(P_{\mathrm{T}} + \sum_{m=1}^{M} P_m\right)$, which is linear and does not violate the convexity. Hence, similar to $\mathcal{P}_3^{\mathrm{WSRM}}$, $\mathcal{P}_3^{\mathrm{WEEM}}$ can also be solved by standard convex optimization methods, which is common in existing works. Solving WEEM problem can be equivalently transformed into finding an ν satisfying $G(\nu) = 0$ [36]. However, this forms a two-fold iterative procedure and each iteration of updating ν requires significant computational complexity. By this method, the complexity of formulating the optimal solution to $\mathcal{P}_2^{\mathrm{WEEM}}$ is intolerable. Nevertheless, the optimal solution of $\mathcal{P}_3^{\mathrm{WEEM}}$ can be obtained via exploiting the KKT conditions and the optimal power allocation of $\mathcal{P}_3^{\mathrm{WSRM}}$ is provided in the following results.

Theorem 3. *$\mathcal{P}_2^{\mathrm{WEEM}}$ is equivalent to $\mathcal{P}_3^{\mathrm{WSRM}}$ if and only if the following condition holds:*

$$C5: \quad P_{\mathrm{T}} \geq \frac{\sum\limits_{m=1}^{M} F_m\left(P_m^{\mathrm{WSR},\star}\right)}{\lambda_{\mathrm{WSR}}} - P_{\max} \equiv P_{\mathrm{T}}^{\mathrm{thr}}, \tag{25}$$

where $\left\{P_m^{\mathrm{WSR},\star}\right\}_{m=1}^{M}$ and λ_{WSR} are the optimal power allocation and Lagrange multiplier at the optimal solution to $\mathcal{P}_3^{\mathrm{WSRM}}$, respectively, $P_{\mathrm{T}}^{\mathrm{thr}}$ represents the constant power threshold.

Proof. See Appendix H. □

Next, we consider the optimal solution to $\mathcal{P}_2^{\mathrm{WEEM}}$. For convenience of statement, we introduce an auxiliary parameter $\chi = \lambda + \nu$, where λ is the Lagrange multiplier for (23a).

Theorem 4. *The optimal solution to $\mathcal{P}_2^{\mathrm{WEEM}}$ is given by*

$$P_m = \max\left\{\frac{\tilde{w}_m(\chi) B_c}{\chi \ln 2} + \frac{\tilde{\rho}_m(\chi)}{\tilde{\xi}_m(\chi)} - \frac{1}{\tilde{\alpha}_m(\chi)\tilde{\xi}_m(\chi)}, \Psi_{1,m}\right\}, \tag{26}$$

where $\tilde{\alpha}_m(\chi)$, $\tilde{w}_m(\chi)$, $\tilde{\xi}_m(\chi)$, $\tilde{\rho}_m(\chi)$ are the same as listed in Theorem 2, and χ is chosen when one of the following conditions holds:

$$C6: \sum_{m=1}^{M} P_m < P_{\max} \text{ and } G(\chi) = 0;$$
$$C7: \sum_{m=1}^{M} P_m = P_{\max} \text{ and } G(\chi) \leq 0.$$

Proof. See Appendix I. □

The optimal power allocation to $\mathcal{P}_2^{\mathrm{WEEM}}$ is fully characterized by Theorem 4. Denote ν and χ of the optimal solution to $\mathcal{P}_2^{\mathrm{WEEM}}$ by $\nu^\star$ and $\chi^\star$, respectively. For $\sum_{m=1}^{M} P_m < P_{\max}$, we have $\lambda = 0$ and $\chi = \nu$ according to (41), implying $\chi > \chi^\star$ if $G(\chi) < 0$ and $\chi \leq \chi^\star$, otherwise. If $\sum_{m=1}^{M} P_m \geq P_{\max}$, we have $\chi \leq \chi^\star$, since P_m is monotonically decreasing in χ according to (26). Thus, the optimal χ can be found via a simple bisection method and the optimal power allocation for the WEEM problem is provided by Algorithm 3. One can observe that (26) is the same as (22), which implies that the power allocation of MG-NOMA WEEM is equal to that of MG-NOMA WSRM if the corresponding Lagrange parameters achieve the same value, i.e., $\chi = \lambda$. It is also consistent with the fact that, given the total transmit power, the WEEM problem can be equivalently transformed into the WSRM problem in MG-NOMA system, since the denominator of the objective in WEEM is fixed.

Algorithm 3 The Optimal Power Allocation for Weighted Energy Efficiency Maximization (OPA WEEM)

```
Obtain {U_{t,m}}_{t=1}^{T_m} using Algorithm 1, ∀m;
Set precision ε_χ > 0, χ_u = max{ dF_1/dΨ_{1,1}, ..., dF_M/dΨ_{1,M} }, χ_d = 0;
while χ_u − χ_d > ε_χ do
    χ = (χ_u + χ_d)/2;
    Obtain P_m according to (26), ∀m;
    if Σ_{m=1}^{M} P_m < P_max and G(χ) ≤ 0 then
        χ_u = χ;
    else
        χ_d = χ;
    end if
end while
Calculate {q*_{n,m}}_{n=1}^{N_m} using Theorem 1, ∀m;
Obtain {p*_{n,m}}_{n=1}^{N_m} according to (4), ∀m.
```

B. Closed-Form Optimal Power Allocation for the Multi-Group WEEM Problem

In this subsection, we derive the optimal power allocation of the MG-NOMA WEEM problem in closed-form. As summarized in Section IV-A, $\mathcal{P}_2^{\mathrm{WEEM}}$ can be solved by updating χ until $G(\chi) = 0$. Then, it poses a new method to establish the power allocation strategy by directly solving the equation $G(\chi) = 0$. To solve the equation, we introduce the Lambert W function [37], termed $W(\cdot)$, which is the inverse function of xe^x. Then, the closed-form optimal solution to $\mathcal{P}_2^{\mathrm{WEEM}}$ is provided in Theorem 5.

Theorem 5. *Suppose that $\chi < \frac{\mathrm{d}F_m}{\mathrm{d}\Lambda_{2,m}}$ holds for each group and $\sum\limits_{m=1}^{M} P_m \leq P_{\max}$, with*

$$\chi = \frac{A}{B} W\left(\frac{B}{A} e^{\frac{C}{A}}\right), A = \sum_{m=1}^{M} \frac{\hat{w}_{1,m} B_c}{\ln 2}, \tag{27a}$$

$$B = P_{\mathrm{T}} + \sum_{m=1}^{M} \frac{\hat{\rho}_{1,m}}{\hat{\xi}_{1,m}} - \frac{1}{\hat{\alpha}_{1,m}\hat{\xi}_{1,m}}, \tag{27b}$$

$$C = \sum_{m=1}^{M} \hat{w}_{1,m} B_c \log\left(\frac{\hat{w}_{1,m}\hat{\alpha}_{1,m}\hat{\xi}_{1,m} B_c}{\ln 2}\right) + \Upsilon_{1,m} - \frac{\hat{w}_{1,m}\hat{\alpha}_{1,m}\hat{\xi}_{1,m} B_c}{\ln 2}, \tag{27c}$$

$$P_m = \frac{\hat{w}_{1,m} B_c}{\chi \ln 2} + \frac{\hat{\rho}_{1,m}}{\hat{\xi}_{1,m}} - \frac{1}{\hat{\alpha}_{1,m}\hat{\xi}_{1,m}}, \forall m. \tag{27d}$$

Then, the optimal solution to $\mathcal{P}_2^{\mathrm{WEEM}}$ is given by (27d).

Proof. See Appendix J. □

This article has been accepted for publication in IEEE Transactions on Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2025.3616144

Algorithm 4 The Closed-Form Optimal Power Allocation for Weighted Energy Efficiency Maximization (CFOPA WEEM)

Dividing the users using Algorithm 1 to obtain $\{\mathcal{U}_{t,m}\}_{t=1}^{T_m}$;
Calculate χ according to (27a), $\forall m$;
Calculate P_m according to (27d), $\forall m$;
Obtain $P_m = \max\{P_m, \Psi_{1,m}\}$, $\forall m$;
Obtain $\{q_{n,m}^\star\}_{n=1}^{N_m}$ using Theorem 1, $\forall m$;
Transform $\{q_{n,m}^\star\}_{n=1}^{N_m}$ into $\{p_{n,m}^\star\}_{n=1}^{N_m}$ according to (4), $\forall m$.

Remark 1. After finding a χ such that $G(\chi) = 0$, P_m shall be achieved by (26) rather than (27d) since $\chi < \frac{\mathrm{d}F_m}{\mathrm{d}\Lambda_{2,m}}$ may not be satisfied for each group and lead to $P_m < \Psi_{1,m}$ with (27d), which violates (23b). Thus, the details of the corresponding solution is presented as Algorithm 4.

C. Complexity Analysis

Now, the complexity of Algorithms 3 and 4 is investigated in this subsection. As mentioned before, the complexity of obtaining the optimal group division scheme and intra-group power allocation is $O(N)$. Assuming that Algorithm 3 converges after I_{WEE} iterations, the inter-group power allocation complexity is $O(I_{\mathrm{WEE}}M)$. Therefore, the overall complexity of Algorithm 3 is $O(N + I_{\mathrm{WEE}}M)$. In Algorithm 4, the inter-group power allocation is obtained in closed form and the related complexity is $O(1)$. The complexity of calculating the parameters in (27b), (27c) and (27d) is $O(N)$. Thus, the total complexity of our proposed closed-form optimal power allocation scheme is $O(N)$.

In this subsection, we also analyze the complexity of several energy efficient power allocation schemes within the prior works. Specifically, the authors in [29] proposed a Dinkelbach algorithm with inner subgradient (D-Subgradient) method. The total complexity of D-Subgradient method is $O(N + C_{\mathrm{DS}}I_{\mathrm{DS}}M)$, where C_{DS} and I_{DS} are the iteration number of the Dinkelbach algorithm and subgradient method, respectively. The Dinkelbach with inner MCPC (D-MCPC) [28] method can be utilized to maximize the weighted energy efficiency, where the fractional programming problem is converted into several subproblems and each subproblem is solved via MCPC method [28]. The total complexity of Dinkelbach with inner MCPC method is $O\left(C_{\mathrm{MCPC}}\left(N^3 + JN^2 + J^2M\right)\right)$, where C_{MCPC} is the number of Dinkelbach iterations and J represents the number of power budget divisions.

V. Numerical Results

In this section, we evaluate the performance of the proposed power allocation methods via numerical simulations. We consider $M = 4$ orthogonal groups with $N_m = 3$ users for each group. The total bandwidth is $B = 4$ MHz and equally divided to M groups with bandwidth $B_c = B/M = 1$ MHz. The users are uniformly distributed in a cell with a radius of 500m and the BS is located at the center. The minimum distance between users is set to be 20m. The channel coefficient $h_{n,m} = PL_{n,m}g_{n,m}$ consists of the path loss exponent $PL_{n,m} = 128.1 + 37.6\log_{10}d_{n,m}$ [38] and the small-scale fading coefficient $g_{n,m}$, where $d_{n,m}$ is the distance in kilometers and $g_{n,m}$ follows independent and identical zero-mean unit-variance complex Gaussian distribution. The noise spectral density is $N_0 = -174$ dBm/Hz and noise power is $\sigma_{n,m}^2 = B_cN_0$. The BS power budget $P_{\max} = 35$ dBm and the circuit power consumption $P_{\mathrm{T}} = 30$ dBm. The weights and QoS requirements are randomly generated to comprehensively cover all potential cases. Thus, the users' weights are generated uniformly at random in $[0, 2]$, ensuring that the mean user weight is equivalent to 1, i.e., $E[w_{n,m}] = 1$. The QoS thresholds are generated uniformly at random in $[0, 2]$ Mbps. We calculate average WSR, WEE, QoS satisfaction ratio and total transmit power of the proposed power allocation method. The QoS satisfaction ratio is the ratio of users which satisfy the QoS constraints (2b) and (3b). In Figs. 1-6, all simulation results are averaged over 1000 random channel profiles.[2]

We compare the proposed methods with several power allocation algorithms, which are designed to address similar problems in MG-NOMA systems within the existing works. The exhaustive search (ES) method is employed to determine the optimal solution. The optimal solution for WSRM and WEEM problems obtained via the ES method are characterized as ES WSRM and ES WEEM, respectively. In the recent RPS [21] scheme, the remaining power is shared among all users after satisfying all constraints. In the Bisection scheme [29], the power allocation is obtained via bisection method. In the D-Subgradient [29] scheme, an energy efficient solution is obtained via outer Dinkelbach method [35] and inner subgradient method. In MCPC [28] scheme, the inter-group power allocation is obtained via dynamic programming [39] and the intra-group power allocation is formulated as the recursive computation. The D-MCPC scheme employs the outer Dinkelbach method [35] and inner MCPC [28] method to solve the MG-NOMA weighted energy efficiency maximization problem.

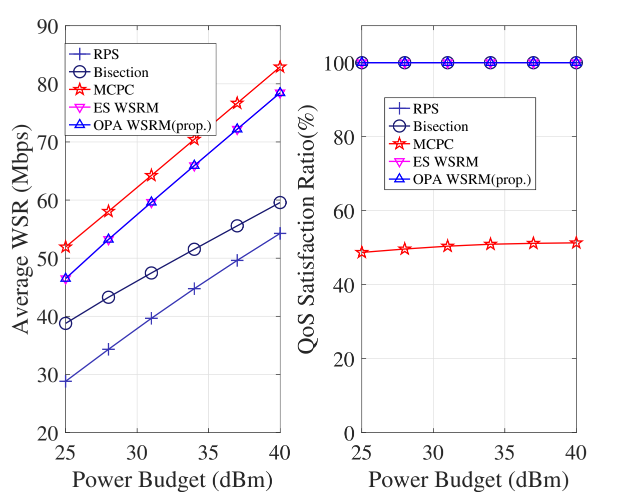


Figure 1. WSR and QoS satisfaction ratio of different algorithms versus power budget.

[2]The source code can be obtained from https://github.com/Slyz060304/Multigroup-NOMA.

This article has been accepted for publication in IEEE Transactions on Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2025.3616144

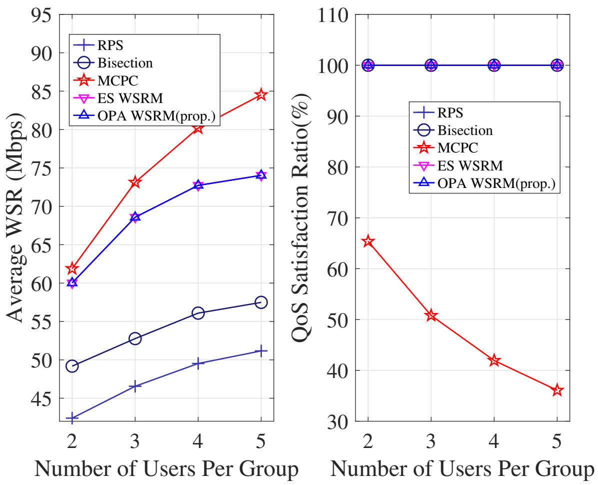


Figure 2. WSR and QoS satisfaction ratio of different algorithms versus the number of users per group.

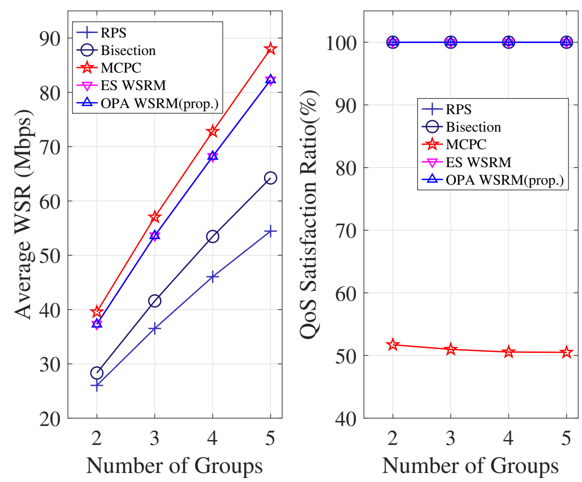


Figure 3. WSR and QoS satisfaction ratio of different algorithms versus the number of groups.

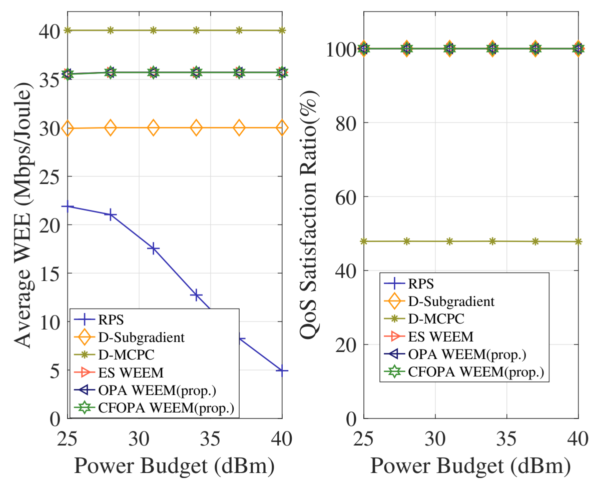


Figure 4. WEE and QoS satisfaction ratio of different algorithms versus power budget.

Fig. 1 displays the average WSR and QoS satisfaction ratio of RPS, Bisection, MCPC, ES WSRM and our proposed OPA WSRM. In Fig. 1, the BS power budget $P_{\max}$ varies from 25 to 40 dBm. The WSR for each algorithm increases with respect to the maximum power since more power budget can be exploited to increase the WSR for each group, as shown in (20). Our proposed OPA WSRM method achieves far better WSR than RPS and Bisection since the user weights are omitted in these algorithms, which makes MG-NOMA system fail to allocate more power to users with higher weights and degrades the system performance. Our proposed method OPA WSRM achieves exactly the same performance with ES WSRM that used the exhaustive search method, demonstrating the optimality of the proposed power allocation strategy. One can observe that the average QoS satisfaction ratio of MCPC is less than 52% because the QoS constraints are neglected in the design of MCPC algorithm and insufficient power allocation is allocated to the users with lower weights.

Fig. 2 shows the performance achieved by various schemes versus the number of users per group. In Fig. 2, the number of users per group N_m is the same among all groups, which varies from 2 to 5. Similar to Fig. 1, Fig. 2 shows that our proposed OPA WSRM achieves better WSR than RPS and Bisection methods. Meanwhile, OPA WSRM displays equivalent performance with the exhaustive search method because of the optimality of our proposed strategy. It is shown that WSR improves as the number of users increases due to the enhanced spectral efficiency resulting from the increasing accessed users. The QoS satisfaction ratio of MCPC is less than 70% because the QoS constraints are omitted with these methods.

The WSR and QoS satisfaction ratio of various schemes versus the number of groups are exhibited in Fig. 3. The bandwidth for each group remains $B_c = 1$ MHz and the corresponding total bandwidth is $B = M * B_c$. In Fig. 3, the number of groups ranges from 2 to 5. Since the user weights are omitted in RPS and Bisection methods, our proposed OPA WSRM achieves superior WSR compared to RPS and Bisection methods. Similar to Fig. 1, it is shown that its QoS satisfaction ratio of MCPC is less than 55%. In Fig. 3, the WSR is proportional to the number of groups and the QoS constraints for all algorithms due to increased users. It is efficient to serve more users by increasing the number of groups. However, more time-frequency resources are also utilized for extra groups.

Figs. 4, 5 and 6 display the average WEE and QoS satisfaction ratio of RPS, D-Subgradient, D-MCPC, ES WEEM, OPA WEEM and CFOPA WEEM versus power budget, number of users per group and number of groups, respectively. Figs. 4, 5 and 6 show that the average WEE of our proposed OPA WEEM and CFOPA WEEM schemeS outperforms RPS and D-Subgradient methods. It is shown in Figs. 4, 5 and 6 that our proposed OPA WEEM method achieves exactly equal performance with the exhaustive search method, which proves that it is the optimal solution to WEEM problem. Figs. 4, 5 and 6 reveal that the QoS satisfaction ratio of D-MCPC remains below 70% due to the absence of QoS constraints in its derivation. Thus, it fails to support more

This article has been accepted for publication in IEEE Transactions on Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2025.3616144

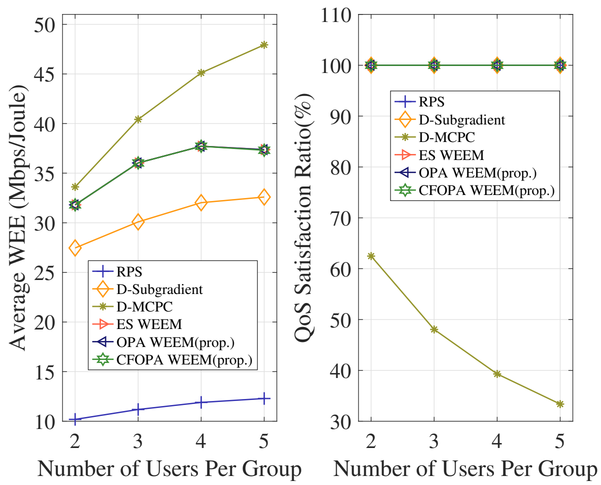


Figure 5. WEE and QoS satisfaction ratio of different algorithms versus the number of users per group.

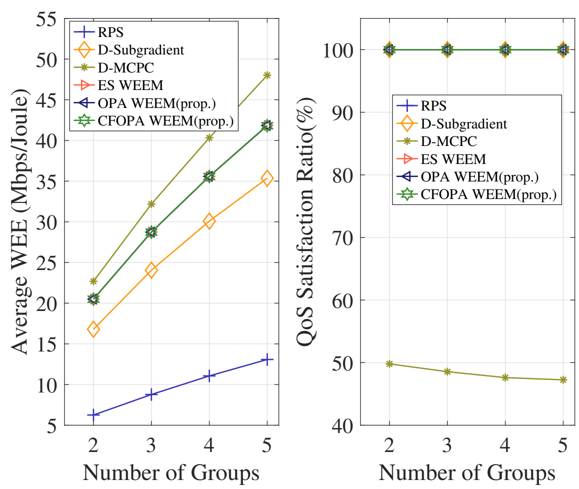


Figure 6. WSR and QoS satisfaction ratio of different algorithms versus the number of groups.

users, although it achieves higher average WEE than our proposed OPA WEEM. It is shown in Fig. 4 that the average WEE of RPS method decreases with power budget because the full power budget is utilized in RPS. In contrast, the average WEE of D-Subgradient, D-MCPC, ES WEEM, OPA WEEM and CFOPA WEEM remains constant with different power budgets, which suggests that the optimal transmit power is not the maximum power budget. As presented in Fig. 5, the WEE improves with the user number because WSR increases with user number even with the same total transmit power. From Fig. 6, one can observe that increasing the number of groups will provide notable spectral efficiency enhancement for all methods. Figs. 4, 5 and 6 show that the average WEE of our proposed CFOPA WEEM is equal to OPA WEEM since the constraint in Theorem 4 will be violated in a few instances and the performance gap can be ignored.

Fig. 7 is obtained in a single channel profile to depict the total transmit power at the optimal point of OPA WEEM with different BS power budgets, where the circuit power consumption $P_{\rm T}$ varies from 30 to 55 dBm. In all cases, total

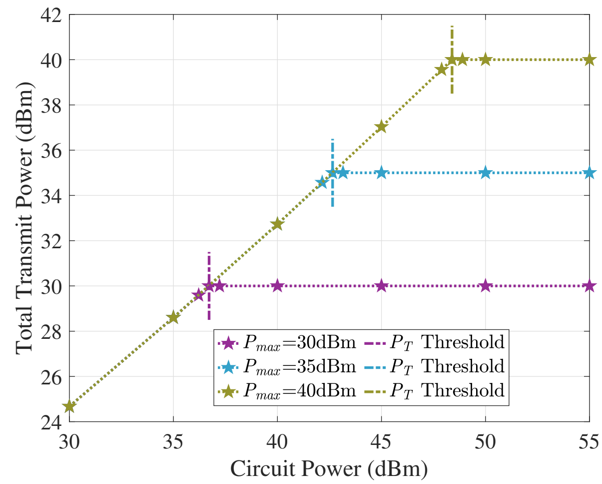


Figure 7. Total transmit power versus circuit power consumption

transmit power is smaller than BS power budget $P_{\max}$ if $P_{\rm T}$ is smaller than the corresponding constant power threshold $P_{\rm T}^{\rm thr}$, which is given by (25). Otherwise, the total transmit power is equal to BS power budget. In this case, the WEEM problem can be equivalent into WSRM problem, which coincides with Theorem 3.

VI. Conclusion

In this paper, we have studied the power allocation in MG-NOMA systems to maximize the weighted sum rate and weighted energy efficiency under QoS constraints. We first provided the closed-form optimal power allocation for SG-NOMA WSRM problem and obtained the WSR function for each group. Next, we have proved that the WSRM and WEEM problems in MG-NOMA systems can be equivalently transformed into inter-group power budget optimization problems. Then, the optimal power allocation for the MG-NOMA WSRM problem was analytically characterized. We have also reformulated the MG-NOMA WEEM problem into an equation of weighted energy efficiency and developed the optimal solution in closed-form. The simulation results have shown that the proposed power allocation schemes achieve superior performance. The extensions of the framework developed in this paper to enable robust designs and under imperfect CSI and SIC are valuable topics for future work.

Appendix

A. Proof of Lemma 1

One can observe that (7) holds for $i = n$. As for $i > n$, $R_{i,m} = r_{i,m}$ is equivalent to $q_{i+1,m} = \eta_{i,m} q_{k,m} - \beta_{i,m}$ for $i = n, \ldots, l-1$. Thus, for $i = n+1, \ldots, l$, $q_{i,m}$ can be expressed by $q_{i-1,m}$ and subsequently by $q_{n,m}$, as indicated in the following result, which completes the proof.

$$
\begin{aligned}
\frac{\mathrm{d}g_{n,l,m}}{\mathrm{d}q_{n,m}} &= \frac{w_{l,m}\alpha_{l,m}\xi_{n,l,m}B_c}{(1+\alpha_{l,m}\xi_{n,l,m}q_{n,m}-\alpha_{l,m}\rho_{n,l,m})\ln 2} - \frac{w_{n-1,m}\alpha_{n-1,m}B_c}{(1+\alpha_{n-1,m}q_{n,m})\ln 2} \\
&= \frac{(w_{l,m}-w_{n-1,m})\,\alpha_{n-1,m}\alpha_{l,m}\xi_{n,l,m}B_c q_{n,m} + (w_{l,m}\alpha_{l,m}\xi_{n,l,m} - w_{n-1,m}\alpha_{n-1,m} + w_{n-1,m}\alpha_{n-1,m}\alpha_{l,m}\rho_{n,l,m})\,B_c}{(1+\alpha_{l,m}\xi_{n,l,m}q_{n,m}-\alpha_{l,m}\rho_{n,l,m})\,(1+\alpha_{n-1,m}q_{n,m})\ln 2} \\
&\geq \frac{(w_{l,m}-w_{n-1,m})\,\alpha_{n-1,m}\alpha_{l,m}\xi_{n,l,m}B_c q_{n,m} + w_{n-1,m}\alpha_{n-1,m}\,(\xi_{n,l,m}-1+\alpha_{l,m}\rho_{n,l,m})\,B_c}{(1+\alpha_{l,m}\xi_{n,l,m}q_{n,m}-\alpha_{l,m}\rho_{n,l,m})\,(1+\alpha_{n-1,m}q_{n,m})\ln 2} \geq 0.
\end{aligned} \tag{29}
$$

$$
\begin{aligned}
\frac{\mathrm{d}^2 g_{n,l,m}}{\mathrm{d}\Gamma_{n,l,m}^2} &= -\frac{w_{l,m}\alpha_{l,m}^2\xi_{n,l,m}^2B_c}{(1+\alpha_{l,m}\xi_{n,l,m}\Gamma_{n,l,m}-\alpha_{l,m}\rho_{n,l,m})^2\ln 2} + \frac{w_{n-1,m}\alpha_{n-1,m}^2B_c}{(1+\alpha_{n-1,m}\Gamma_{n,l,m})^2\ln 2} \\
&= \frac{w_{n-1,m}\,(w_{l,m}-w_{n-1,m})\,\alpha_{n-1,m}^2B_c}{w_{l,m}\,(1+\alpha_{n-1,m}\Gamma_{n,l,m})^2\ln 2} < 0.
\end{aligned} \tag{30}
$$

$$
\begin{aligned}
q_{n,m}^{\star} &= \xi_{J_{t,m},n,m}q_{J_{t,m},m}^{\star} - \rho_{J_{t,m},n,m} \\
&= \min\left\{\xi_{J_{t,m},n,m}\left(\xi_{1,J_{t,m},m}P_m - \rho_{1,J_{t,m},m}\right) - \rho_{J_{t,m},n,m}, \xi_{J_{t,m},n,m}\Phi_{t,m} - \rho_{J_{t,m},n,m}\right\} \\
&= \min\left\{\xi_{1,n,m}P_m - \rho_{1,n,m}, \xi_{J_{t,m},n,m}\Phi_{t,m} - \rho_{J_{t,m},n,m}\right\}.
\end{aligned} \tag{31}
$$

$$
\begin{aligned}
q_{i,m} &= \eta_{i-1,m}q_{i-1,m} - \beta_{i-1,m} \\
&= \eta_{i-2,m}\eta_{i-1,m}q_{i-2,m} - \eta_{i-1,m}\beta_{i-2,m} - \beta_{i-1,m} \\
&= \cdots = \left(\prod_{j=n}^{i-1}\eta_{j,m}\right) q_{n,m} - \sum_{j=n}^{i-1}\xi_{j+1,i,m}\beta_{j,m} \\
&= \xi_{n,i,m}q_{n,m} - \rho_{n,i,m}.
\end{aligned} \tag{28}
$$

B. Proof of Proposition 1

$R_{i,m} \geq r_{i,m}$ is equivalent to $q_{i+1,m} \leq \eta_{i,m}q_{i,m} - \beta_{i,m}$ for $i = n, \ldots, l-1$. Through similar iterative substitutions in (28), we have $q_{l,m} \leq \xi_{n,l,m}q_{n,m} - \rho_{n,l,m}$. By substituting (6c) into it with $l = N_m$, we have $\xi_{n,N_m,m}q_{n,m} - \rho_{n,N_m,m} \geq q_{N_m,m} \geq \theta_{N_m,m}$ and further $q_{n,m} \geq \Psi_{n,m}$ if $R_{i,m} \geq r_{i,m}$ for $i = n, \ldots, N_m$, which also holds for $n = 1$. Thus, we have $P_m \geq \Psi_{1,m}$ since $P_m = q_{1,m}$ and $\mathcal{P}_{2,m}^{\mathrm{WSRM}}$ is feasible only if $P_m \geq \Psi_{1,m}$. Besides, the solution with $R_{i,m} = r_{i,m}, \forall i$, is feasible and $P_m = \Psi_{1,m}$ since $\xi_{1,N_m,m}q_{1,m} - \rho_{1,N_m,m} = q_{N_m,m} = \theta_{N_m,m}$ according to (28). Thus, $\mathcal{P}_{2,m}^{\mathrm{WSRM}}$ is feasible if $P_m \geq \Psi_{1,m}$, which completes the proof.

C. Proof of Lemma 2

Given $w_{n-1,m} \leq w_{l,m}$, the derivative of $g_{n,l,m}(q_{n,m})$ is given by (29) on the top of next page. Consequently, $g_{n,l,m}(q_{n,m})$ is a nondecreasing function and the optimum $q_{n,m}$ is achieved by maximizing $q_{n,m}$ until the feasibility constraint $\eta_{n-1,m}q_{n-1,m} - \beta_{n-1,m} \geq q_{n,m}$ takes the equality.

D. Proof of Lemma 3

Assuming C3 holds, by setting the first-order derivative of $g_{n,l,m}(q_{n,m})$ to zero, we obtain a unique root $\Gamma_{n,l,m}$ according to (29), where $\frac{\mathrm{d}g_{n,l,m}}{\mathrm{d}\Gamma_{n,l,m}} = 0$. Moreover, it follows from (30) on the top of next page that the second order derivative of $g_{n,l,m}(q_{n,m})$ at $\Gamma_{n,l,m}$ is $\frac{\mathrm{d}^2 g_{n,l,m}}{\mathrm{d}\Gamma_{n,l,m}^2} < 0$, corroborating that $\Gamma_{n,l,m}$ is the maximizer of $g_{n,l,m}(q_{n,m})$. As proved in Appendix B, $q_{n,m} \geq \Psi_{n,m}$ is required due to the feasibility constraints and we have $\Omega_{n,l,m} = \max\{\Gamma_{n,l,m}, \Psi_{n,m}\}$.

E. Proof of Lemma 4

Lemma 4 is proved via contradiction. For any feasible solution $\xi_{n,l+1,m}\Omega_{n,l,m} - \rho_{n,l+1,m} \leq \Omega_{l+1,k,m}$ with $q_{l+1,m} > \Omega_{l+1,k,m}$, the solution with $\tilde{q}_{n,m} = (q_{l+1,m} + \rho_{n,l+1,m})/\xi_{n,l+1,m}$ yields higher weighted sum rate, since

$$
\begin{aligned}
&\xi_{n,l+1,m}q_{n,m} - \rho_{n,l+1,m} \\
&\quad \geq q_{l+1,m} = \xi_{n,l+1,m}\tilde{q}_{n,m} - \rho_{n,l+1,m} \\
&\quad > \Omega_{l+1,k,m} \geq \xi_{n,l+1,m}\Omega_{n,l,m} - \rho_{n,l+1,m},
\end{aligned} \tag{32}
$$

and thus $q_{n,m} \geq \tilde{q}_{n,m} > \Omega_{n,l,m}$. As for

$$
\Omega_{l+1,k,m} \geq \xi_{n,l+1,m}q_{n,m} - \rho_{n,l+1,m} \geq q_{l+1,m},
$$

the solution with $\tilde{q}_{l+1,m} = \xi_{n,l+1,m}q_{n,m} - \rho_{n,l+1,m}$ achieves an improvement since $\Omega_{l+1,k,m} \geq \tilde{q}_{l+1,m} \geq q_{l+1,m}$. If

$$
\xi_{n,l+1,m}q_{n,m} - \rho_{n,l+1,m} \geq \Omega_{l+1,k,m} \geq q_{l+1,m},
$$

the solution with $\tilde{q}_{l+1,m} = \Omega_{l+1,k,m}$ and $\tilde{q}_{n,m} = (\Omega_{l+1,k,m} + \rho_{n,l+1,m})/\xi_{n,l+1,m}$ delivers a higher weighted sum rate since $q_{n,m} \geq \tilde{q}_{n,m} \geq \Omega_{n,l,m}$ with

$$
\begin{aligned}
&\xi_{n,l+1,m}q_{n,m} - \rho_{n,l+1,m} \\
&\quad \geq \Omega_{l+1,k,m} \geq \xi_{n,l+1,m}\Omega_{n,l,m} - \rho_{n,l+1,m}.
\end{aligned} \tag{33}
$$

Thus, we can increase the weighted sum rate with $q_{l+1,m} = \xi_{n,l+1,m}q_{n,m} - \rho_{n,l+1,m}$ if C4 is violated. Consequently, the optimal solution achieves $R_{l,m}^{\star} = r_{l,m}$ with violation of C4.

F. Proof of Theorem 1

As mentioned above, we have $q_{1,m} = P_m$ according to (6a). Consequently, within the integrated subgroup $\mathcal{U}_{1,m}$, it can be verified that $q_{n,m}^{\star} = \xi_{1,n,m}P_m - \rho_{1,n,m}$. It follows from (11)

This article has been accepted for publication in IEEE Transactions on Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2025.3616144

that $\Phi_{t,m}$ is the maximizer of $\mathcal{U}_{t,m}$ provided that C3 is satisfied for $t > 1$. The optimal solution achieves $q^{\star}_{J_{t,m},m} = \Phi_{t,m}$ if P_m is sufficiently large, taking into account that $\Phi_{t,m}$ is the maximizer of $\mathcal{U}_{t,m}$ and the feasibility constraints are satisfied due to $\psi_{t,m}\Phi_{t,m} - \delta_{t,m} \geq \Phi_{t+1,m}$. However, if P_m is not sufficiently large, the optimum reduces to the feasible upper bound, represented as

$$q^{\star}_{J_{t,m},m} = \xi_{1,J_{t,m},m}P_m - \rho_{1,J_{t,m},m},$$

since the weighted sum rate function $g_{J_{t,m},L_{t,m},m}\left(q_{J_{t,m},m}\right)$ is monotonically nondecreasing in $q_{J_{t,m},m}$ if $q_{J_{t,m},m} < \Phi_{t,m}$. Thus, we have

$$q^{\star}_{J_{t,m},m} = \min\left\{\xi_{1,J_{t,m},m}P_m - \rho_{1,J_{t,m},m}, \Phi_{t,m}\right\}, \tag{34}$$

for $t > 1$. By substituting (28) into (34), $q^{\star}_{n,m}$ can further be formulated as (31) for $n \in \mathcal{U}_{t,m}$, $t > 1$.

G. Proof of Theorem 2

The Lagrange of $\mathcal{P}_3^{\mathrm{WSRM}}$ can be written as

$$L = \sum_{m=1}^{M} F_m\left(P_m\right) + \lambda\left(P_{\max} - \sum_{m=1}^{M} P_m\right) + \sum_{m=1}^{M} \mu_m\left(P_m - \Psi_{1,m}\right), \tag{35}$$

where λ and $\{\mu_m\}_{m=1}^{M}$ are Lagrange multipliers, and the KKT conditions are satisfied:

$$\frac{\partial L}{\partial P_m} = \frac{\mathrm{d}F_m}{\mathrm{d}P_m} - \lambda + \mu_m = 0, \forall m, \tag{36}$$

$$\lambda\left(P_{\max} - \sum_{m=1}^{M} P_m\right) = 0, \lambda \geq 0, \tag{37}$$

$$\mu_m\left(P_m - \Psi_{1,m}\right) = 0, \mu_m \geq 0, \forall m. \tag{38}$$

If $\mu_m > 0$, we have $P_m = \Psi_{1,m}$ and $\lambda > \frac{\mathrm{d}F_m}{\mathrm{d}\Psi_{1,m}}$. If $\mu_m = 0$, we have $\lambda = \frac{\mathrm{d}F_m}{\mathrm{d}P_m} \leq \frac{\mathrm{d}F_m}{\mathrm{d}\Psi_{1,m}}$ and $P_m \geq \Psi_{1,m}$ since $\frac{\mathrm{d}F_m}{\mathrm{d}P_m}$ is monotonically decreasing in P_m, where $\frac{\mathrm{d}F_m}{\mathrm{d}\Psi_{1,m}}$ is the derivative of $F_m\left(P_m\right)$ at $\Psi_{1,m}$ and can be obtained via (20). Given the derivative value λ, it can be mapped to an associated group power P_m, and the coefficients $\tilde{\alpha}_m\left(\lambda\right)$, $\tilde{w}_m\left(\lambda\right)$, $\tilde{\xi}_m\left(\lambda\right)$ and $\tilde{\rho}_m\left(\lambda\right)$ are contingent upon the intervals to which λ corresponds. Then, it follows from (36) that $\lambda = \frac{\mathrm{d}F_m}{\mathrm{d}P_m} + \mu_m > 0$. Finally, from (37), we have $\sum_{m=1}^{M} P_m = P_{\max}$, which completes the proof.

H. Proof of Theorem 3

The Lagrange of $\mathcal{P}_3^{\mathrm{WEEM}}$ can be written as

$$L = \sum_{m=1}^{M} F_m\left(P_m\right) - \nu\left(P_{\mathrm{T}} + \sum_{m=1}^{M} P_m\right) + \lambda\left(P_{\max} - \sum_{m=1}^{M} P_m\right) + \sum_{m=1}^{M} \mu_m\left(P_m - \Psi_{1,m}\right), \tag{39}$$

where λ and $\{\mu_m\}_{m=1}^{M}$ are Lagrange multipliers. Since $F_m\left(P_m\right)$ is concave and the constraints are linear, $\mathcal{P}_3^{\mathrm{WEEM}}$ is a convex problem and the optimal solution is characterized by the KKT conditions:

$$\frac{\partial L}{\partial P_m} = \frac{\mathrm{d}F_m}{\mathrm{d}P_m} - \nu - \lambda + \mu_m = 0, \forall m, \tag{40}$$

$$\lambda\left(P_{\max} - \sum_{m=1}^{M} P_m\right) = 0, \lambda \geq 0, \tag{41}$$

$$\mu_m\left(P_m - \Psi_{1,m}\right) = 0, \mu_m \geq 0, \forall m. \tag{42}$$

We denote ν at the optimal value of $\mathcal{P}_2^{\mathrm{WEEM}}$ as $\nu^{\star}$ and λ at the optimal point of $\mathcal{P}_2^{\mathrm{WEEM}}$ as $\lambda^{\star}$ when $\nu = \nu^{\star}$. Suppose that $\mathcal{P}_2^{\mathrm{WEEM}}$ is equivalent to $\mathcal{P}_3^{\mathrm{WSRM}}$, we can apply Algorithm 2 to solve $\mathcal{P}_2^{\mathrm{WEEM}}$ and achieve exactly the same solution, i.e., $P_m^{\star} = P_m^{\mathrm{WSR},\star}$. It follows from (36) and (40) that $\lambda_{\mathrm{WSR}} = \nu^{\star} + \lambda^{\star}$. Then, from [40], we have $G\left(\lambda_{\mathrm{WSR}}\right) \leq 0$ if $\lambda_{\mathrm{WSR}} \geq \nu^{\star}$. In contrast, $G\left(\lambda_{\mathrm{WSR}}\right) > 0$ suggests $\lambda_{\mathrm{WSR}} < \nu^{\star} \leq \nu^{\star} + \lambda^{\star}$, which means the equivalence between $\mathcal{P}_2^{\mathrm{WEEM}}$ and $\mathcal{P}_3^{\mathrm{WSRM}}$ does not hold. Thereby, the validity of the transformation is equivalent to $G\left(\lambda_{\mathrm{WSR}}\right) \leq 0$. If $G\left(\lambda_{\mathrm{WSR}}\right) \leq 0$, $\mathcal{P}_2^{\mathrm{WEEM}}$ and $\mathcal{P}_3^{\mathrm{WSRM}}$ yield the same optimal solution, thus $P_m = P_m^{\mathrm{WSR},\star}$ and $\sum_{m=1}^{M} P_m = P_{\max}$. Subsequently, $G\left(\lambda_{\mathrm{WSR}}\right) \leq 0$ can be transformed into $\sum\limits_{m=1}^{M} F_m\left(P_m^{\mathrm{WSR},\star}\right) - \lambda_{\mathrm{WSR}}\left(P_{\mathrm{T}} + P_{\max}\right) \leq 0$, thereby proving Theorem 3.

I. Proof of Theorem 4

If $P_{\mathrm{T}} < P_{\mathrm{T}}^{\mathrm{thr}}$, the WEEM can not be equivalently transformed into WSRM and we have $\sum_{m=1}^{M} P_m < P_{\max}$. It follows from (41) that $\lambda = 0$, $\chi = \nu$. Next, from (40) and (42), it can be verified that $P_m = \Psi_{1,m}$, $\mu_m > 0$ if $\chi > \frac{\mathrm{d}F_m}{\mathrm{d}\Psi_{1,m}}$ and $\chi = \frac{\mathrm{d}F_m}{\mathrm{d}P_m}$ if $P_m > \Psi_{1,m}$. Then, the power budget P_m can be obtained via (26) and we have $G\left(\nu\right) = G\left(\chi\right) = 0$. As for $P_{\mathrm{T}} \geq P_{\mathrm{T}}^{\mathrm{thr}}$, the WEEM can be equivalently transformed into WSRM and we have $\chi = \lambda_{\mathrm{WSR}}$ since (26) is the same as that of (22). In this case, we have

$$\begin{aligned} G\left(\chi\right) &= \sum_{m=1}^{M} F_m\left(P_m\right) - \left(\nu + \lambda\right)\left(P_{\mathrm{T}} + \sum_{m=1}^{M} P_m\right) \\ &= G\left(\nu\right) - \lambda\left(P_{\mathrm{T}} + \sum_{m=1}^{M} P_m\right) \\ &= -\lambda\left(P_{\mathrm{T}} + \sum_{m=1}^{M} P_m\right) \leq 0, \end{aligned}$$

which completes the proof.

J. Proof of Theorem 5

Suppose that $\chi < \frac{\mathrm{d}F_m}{\mathrm{d}\Lambda_{2,m}}$, we have $P_m > \Lambda_{2,m} \geq \Psi_{1,m}$ since P_m is monotonically decreasing in χ. Consequently, it can be verified that $\mu_m = 0$ according to (42). Then, $\lambda = 0$

This article has been accepted for publication in IEEE Transactions on Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2025.3616144

according to (41) and $\sum_{m=1}^{M} P_m < P_{\max}$. In this case, $\frac{\mathrm{d}F_m}{\mathrm{d}P_m} = \chi$ and P_m is specified in (27d). Next, we expand $G(\chi)$ as

$$\begin{aligned} G(\chi) &= \sum_{m=1}^{M} F_m(P_m) - \chi\left(P_{\mathrm{T}} + \sum_{m=1}^{M} P_m\right) \\ &= \sum_{m=1}^{M} \left[G_{1,m}(P_m) + \Upsilon_{1,m}\right] - \chi\left(P_{\mathrm{T}} + \sum_{m=1}^{M} P_m\right) \\ &= \sum_{m=1}^{M} \left[\hat{w}_{1,m} B_c \log\left(\frac{\hat{w}_{1,m}\hat{\alpha}_{1,m}\hat{\xi}_{1,m} B_c}{\nu \ln 2}\right) + \Upsilon_{1,m}\right] \\ &\quad -\chi\left(P_{\mathrm{T}} + \sum_{m=1}^{M} \frac{\hat{w}_{1,m} B_c}{\nu \ln 2} + \frac{\hat{\rho}_{1,m}}{\hat{\xi}_{1,m}} - \frac{1}{\hat{\alpha}_{1,m}\hat{\xi}_{1,m}}\right) \\ &= C - A\ln(\chi) - B\nu. \end{aligned} \tag{43}$$

It can be verified that $\chi = \frac{A}{B} W\left(\frac{B}{A} e^{\frac{C}{A}}\right)$ is equal to $\frac{B\nu}{A} = W\left(\frac{B}{A} e^{\frac{C}{A}}\right)$, which implies $\frac{B\chi}{A} e^{\frac{B\chi}{A}} = \frac{B}{A} e^{\frac{C}{A}}$ since $W(\cdot)$ is the inverse function of xe^x. So $\ln(\chi) = \frac{C - B\chi}{A}$ and thus $A\ln(\chi) + B\chi - C = 0$. Therefore, we have $G(\chi) = 0$ and $\chi^\star = \frac{A}{B} W\left(\frac{B}{A} e^{\frac{C}{A}}\right)$.

References

[1] W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, "The road towards 6G: A comprehensive survey," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 334–366, Feb. 2021.

[2] C.-X. Wang, X. You, X. Gao, X. Zhu, Z. Li, C. Zhang, H. Wang, Y. Huang, Y. Chen, H. Haas, J. S. Thompson, E. G. Larsson, M. D. Renzo, W. Tong, P. Zhu, X. Shen, H. V. Poor, and L. Hanzo, "On the road to 6G: Visions, requirements, key technologies, and testbeds," *IEEE Commun. Surv. Tutorials*, vol. 25, no. 2, pp. 905–974, Feb. 2023.

[3] F. Guo, F. R. Yu, H. Zhang, X. Li, H. Ji, and V. C. M. Leung, "Enabling massive IoT toward 6G: A comprehensive survey," *IEEE Internet Things J.*, vol. 8, no. 15, pp. 11 891–11 915, Mar. 2021.

[4] Y. Huang, C. Zhang, J. Wang, Y. Jing, L. Yang, and X. You, "Signal processing for MIMO-NOMA: Present and future challenges," *IEEE Wireless Commun.*, vol. 25, no. 2, pp. 32–38, Apr. 2018.

[5] X. Chen, D. W. K. Ng, W. Yu, E. G. Larsson, N. Al-Dhahir, and R. Schober, "Massive access for 5G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 3, pp. 615–637, Sep. 2021.

[6] Y. Liu, S. Zhang, X. Mu, Z. Ding, R. Schober, N. Al-Dhahir, E. Hossain, and X. Shen, "Evolution of NOMA toward next generation multiple access (NGMA) for 6G," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 4, pp. 1037–1071, Jan. 2022.

[7] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, "A survey of non-orthogonal multiple access for 5G," *IEEE Commun. Surv. Tutorials*, vol. 20, no. 3, pp. 2294–2323, May 2018.

[8] S. M. 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. Surv. Tutorials*, vol. 19, no. 2, pp. 721–742, Oct. 2017.

[9] S. Timotheou and I. Krikidis, "Fairness for non-orthogonal multiple access in 5G systems," *IEEE Signal Process Lett.*, vol. 22, no. 10, pp. 1647–1651, Mar. 2015.

[10] K. Li, M. Cui, G. Zhang, Q. Wu, D. Li, and J. He, "Robust transmit and reflect beamforming design for IRS-assisted offshore NOMA communication systems," *IEEE Trans. Veh. Technol.*, vol. 73, no. 1, pp. 783–798, Jan. 2024.

[11] Y. Wang, J. Wang, D. W. Kwan Ng, R. Schober, and X. Gao, "A minimum error probability NOMA design," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4221–4237, Feb. 2021.

[12] R. Liu, K. Guo, K. An, S. Zhu, C. Li, and L. Gao, "Performance evaluation of NOMA-based cognitive integrated satellite terrestrial relay networks with primary interference," *IEEE Access*, vol. 9, pp. 71 422–71 434, May 2021.

[13] J. Wang, Q. Peng, Y. Huang, H.-M. Wang, and X. You, "Convexity of weighted sum rate maximization in NOMA systems," *IEEE Signal Process Lett.*, vol. 24, no. 9, pp. 1323–1327, Jul. 2017.

[14] I. Abu Mahady, E. Bedeer, S. Ikki, and H. Yanikomeroglu, "Sum-rate maximization of NOMA systems under imperfect successive interference cancellation," *IEEE Commun. Lett.*, vol. 23, no. 3, pp. 474–477, Mar. 2019.

[15] T. Wang, F. Fang, and Z. Ding, "An SCA and relaxation based energy efficiency optimization for multi-user RIS-assisted NOMA networks," *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 6843–6847, Mar. 2022.

[16] W. U. Khan, M. A. Javed, T. N. Nguyen, S. Khan, and B. M. Elhalawany, "Energy-efficient resource allocation for 6G backscatter-enabled NOMA IoV networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 9775–9785, Sep. 2022.

[17] J. Zhu, Y. Huang, J. Wang, K. Navaie, and Z. Ding, "Power efficient IRS-assisted NOMA," *IEEE Trans. Commun.*, vol. 69, no. 2, pp. 900–913, Oct. 2021.

[18] B. Xu, Z. Xiang, P. Ren, and X. Guo, "Outage performance of downlink full-duplex network-coded cooperative NOMA," *IEEE Wireless Commun. Lett.*, vol. 10, no. 1, pp. 26–29, Jan. 2021.

[19] 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.

[20] M. Vaezi, A. Azari, S. R. Khosravirad, M. Shirvanimoghaddam, M. M. Azari, D. Chasaki, and P. Popovski, "Cellular, wide-area, and non-terrestrial IoT: A survey on 5G advances and the road toward 6G," *IEEE Commun. Surv. Tutorials*, vol. 24, no. 2, pp. 1117–1174, Feb. 2022.

[21] M. Pavan Reddy, "Enhanced resource management for RIS-assisted NOMA: A fractional programming approach," *IEEE Commun. Lett.*, early access, Jun., 9, 2025.

[22] X. Wu, Y. Ko, and A. M. Tyrrell, "Distributed multi-agent reinforcement learning for heterogeneous NOMA-ALOHA systems," *IEEE Trans. Cognit. Commun. Networking*, vol. 11, no. 3, pp. 1902–1912, Jun. 2025.

[23] S. Solaiman, "Optimizing subchannel assignment and power allocation for network slicing in high-density NOMA networks: A Q-learning approach," *IEEE Access*, vol. 13, pp. 24 323–24 335, Feb. 2025.

[24] H. Zhang, H. Zhang, K. Long, and G. K. Karagiannidis, "Deep learning based radio resource management in NOMA networks: User association, subchannel and power allocation," *IEEE Trans. Network Sci. Eng.*, vol. 7, no. 4, pp. 2406–2415, Jun. 2020.

[25] D. Kim, H. Jung, I.-H. Lee, and D. Niyato, "Novel resource allocation algorithm for IoT networks with multicarrier NOMA," *IEEE Internet Things J.*, vol. 11, no. 18, pp. 30 354–30 367, Sept. 2024.

[26] H. Zhu, Q. Wu, X.-J. Wu, Q. Fan, P. Fan, and J. Wang, "Decentralized power allocation for MIMO-NOMA vehicular edge computing based on deep reinforcement learning," *IEEE Internet Things J.*, vol. 9, no. 14, pp. 12 770–12 782, Jul. 2022.

[27] 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, Jan. 2017.

[28] L. Salaün, M. Coupechoux, and C. S. Chen, "Joint subcarrier and power allocation in NOMA: Optimal and approximate algorithms," *IEEE Trans. Signal Process.*, vol. 68, pp. 2215–2230, Mar. 2020.

[29] S. Rezvani, E. A. Jorswieck, R. Joda, and H. Yanikomeroglu, "Optimal power allocation in downlink multicarrier NOMA systems: Theory and fast algorithms," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 4, pp. 1162–1189, Apr. 2022.

[30] F. Fang, J. Cheng, and Z. Ding, "Joint energy efficient subchannel and power optimization for a downlink NOMA heterogeneous network," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1351–1364, Nov. 2019.

[31] Z. Xiao, L. Zhu, Z. Gao, D. O. Wu, and X.-G. Xia, "User fairness non-orthogonal multiple access (NOMA) for millimeter-wave communications with analog beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 7, pp. 3411–3423, May 2019.

[32] J. Choi, "Power allocation for max-sum rate and max-min rate proportional fairness in NOMA," *IEEE Commun. Lett.*, vol. 20, no. 10, pp. 2055–2058, Oct. 2016.

[33] K. G. Binmore, *Mathematical Analysis: A straightforward approach*. Cambridge University Press, 1982.

[34] D. Adionel Guimaraes, G. H. Faria Floriano, and L. Silvestre Chaves, "A tutorial on the CVX system for modeling and solving convex optimization problems," *IEEE Lat. Am. Trans.*, vol. 13, no. 5, pp. 1228–1257, May 2015.

[35] W. Dinkelbach, "On nonlinear fractional programming," *Manage. Sci.*, vol. 13, no. 7, pp. 492–498, Mar. 1967.

[36] R. G. Ródenas, M. L. López, and D. Verastegui, "Extensions of Dinkelbach's algorithm for solving non-linear fractional programming problems," *Top*, vol. 7, no. 1, pp. 33–70, Jun. 1999.

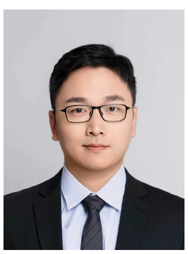
[37] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert W function," *Adv. Comput. Math.*, vol. 5, pp. 329–359, Dec. 1996.

[38] *3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects(Release 9)*, 3GPP TR 36.814 V9.2.0 (2017-03) Std., 2017.

[39] H. Kellerer, U. Pferschy, and D. Pisinger, *Knapsack Problems*. Berlin, Heidelberg: Springer, 2004.

[40] C. Isheden, Z. Chong, E. Jorswieck, and G. Fettweis, "Framework for link-level energy efficiency optimization with informed transmitter," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2946–2957, Aug. 2012.

Qian Peng received the B.E. and M.S. degrees from Southeast University, Nanjing, China, in 2016 and 2019, respectively. He is currently pursuing the Ph.D. degree in information and communication engineering with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing. His research interests include next-generation multiple access and optimization theory for wireless communications.



Rui Zhou received the B.Eng. degree in information engineering from Southeast University, Nanjing, China, in 2017, and the Ph.D. degree from the Hong Kong University of Science and Technology (HKUST), Hong Kong, in 2021. He is currently a Research Scientist at the Shenzhen Research Institute of Big Data and an Adjunct Assistant Professor at the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China. His research interests include optimization algorithms, statistical signal processing, machine learning, and financial engineering.

Jiaheng Wang (M'10–SM'14) received the Ph.D. degree in electronic and computer engineering from the Hong Kong University of Science and Technology, Kowloon, Hong Kong, in 2010, and the B.E. and M.S. degrees from the Southeast University, Nanjing, China, in 2001 and 2006, respectively.

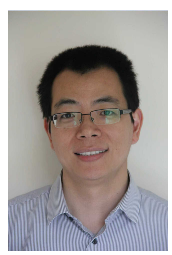
He is currently a Full Professor at the National Mobile Communications Research Laboratory (NCRL), Southeast University, Nanjing, China. He is also with Purple Mountain Laboratories Nanjing, China. From 2010 to 2011, he was with the Signal Processing Laboratory, KTH Royal Institute of Technology, Stockholm, Sweden. He also held visiting positions at the Friedrich Alexander University Erlangen-Nürnberg, Nürnberg, Germany, and the University of Macau, Macau. His research interests are mainly on communication systems, wireless networks and network security.

Dr. Wang has published more than 200 articles on international journals and conferences. He serves as an Editor for the IEEE Transactions on Wireless Communications and an Editor for the IEEE Transactions on Communications. He was a Senior Area Editor for the IEEE Signal Processing Letters. He was a recipient of the Humboldt Fellowship for Experienced Researchers and the best paper awards of IEEE GLOBECOM 2019, ADHOCNETS 2019, and WCSP 2022 and 2014.

Xiang-Gen Xia (M'97, S'00, F'09) received his B.S. degree in mathematics from Nanjing Normal University, Nanjing, China, and his M.S. degree in mathematics from Nankai University, Tianjin, China, and his Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 1983, 1986, and 1992, respectively.

He was a Senior/Research Staff Member at Hughes Research Laboratories, Malibu, California, during 1995-1996. In September 1996, he joined the Department of Electrical and Computer Engineering, University of Delaware, Newark, Delaware, where he is the Charles Black Evans Professor. His current research interests include space-time coding, MIMO and OFDM systems, digital signal processing, and SAR and ISAR imaging. Dr. Xia is the author of the book Modulated Coding for Intersymbol Interference Channels (New York, Marcel Dekker, 2000), and a co-author of the book Array Beamforming Enabled Wireless Communications (New York, CRC Press, 2023).

Dr. Xia received the National Science Foundation (NSF) Faculty Early Career Development (CAREER) Program Award in 1997, the Office of Naval Research (ONR) Young Investigator Award in 1998, and the Outstanding Overseas Young Investigator Award from the National Nature Science Foundation of China in 2001. He received the 2019 Information Theory Outstanding Overseas Chinese Scientist Award, The Information Theory Society of Chinese Institute of Electronics. Dr. Xia has served as an Associate Editor for numerous international journals including IEEE Transactions on Signal Processing, IEEE Transactions on Wireless Communications, IEEE Transactions on Mobile Computing, and IEEE Transactions on Vehicular Technology. Dr. Xia is Technical Program Chair of the Signal Processing Symp., Globecom 2007 in Washington D.C. and the General Co-Chair of ICASSP 2005 in Philadelphia.



Zhiguo Ding (S'03-M'05-F'20) received his B.Eng from the Beijing University of Posts and Telecommunications in 2000, and the Ph.D degree from Imperial College London in 2005. He is currently a Professor in Communications at Khalifa University, and has also been affiliated with the University of Manchester and Princeton University.

Dr Ding' research interests are 6G networks, multiple access, energy harvesting networks and statistical signal processing. He is serving as an Area Editor for the *IEEE Transactions on Wireless Communications*, and *IEEE Open Journal of the Communications Society*, an Editor for *IEEE Transactions on Vehicular Technology*, and was an Editor for *IEEE Wireless Communication Letters*, *IEEE Transactions on Communications*, *IEEE Communication Letters* from 2013 to 2016. He recently received the EU Marie Curie Fellowship 2012-2014, the Top IEEE TVT Editor 2017, IEEE Heinrich Hertz Award 2018, IEEE Jack Neubauer Memorial Award 2018, IEEE Best Signal Processing Letter Award 2018, Friedrich Wilhelm Bessel Research Award 2020, IEEE SPCC Technical Recognition Award 2021 and IEEE VTS Best Magazine Paper Award 2023. He is a Fellow of the IEEE, a Distinguished Lecturer of IEEE ComSoc, and a Web of Science Highly Cited Researcher in two categories 2022.