

Joint Beam Management and Power Allocation in THz-NOMA Networks

Zhiguo Ding, *Fellow, IEEE* and H. Vincent Poor, *Life Fellow, IEEE*

Abstract

This paper investigates how to apply non-orthogonal multiple access (NOMA) as an add-on in terahertz (THz) networks. In particular, prior to the implementation of NOMA, it is assumed that there exists a legacy THz system, where spatial beams have already been configured to serve legacy primary users. The aim of this paper is to study how these pre-configured spatial beams can be used as a type of bandwidth resources, on which additional secondary users are served without degrading the performance of the legacy primary users. A joint beam management and power allocation problem is first formulated as a mixed combinatorial non-convex optimization problem, and then solved by two methods with different performance-complexity tradeoffs, one based on the branch and bound method and the other based on successive convex approximation. Both analytical and simulation results are presented to illustrate the new features of beam-based resource allocation in THz-NOMA networks and also demonstrate that those pre-configured spatial beams can be employed to improve the system throughput and connectivity in a spectrally efficient manner.

Index Terms

Non-orthogonal multiple access (NOMA), Terahertz (THz), the branch and bound method, successive convex approximation, beam management, power allocation.

I. INTRODUCTION

Terahertz (THz) communications and non-orthogonal multiple access (NOMA) are two key enabling technologies for the envisioned sixth generation (6G) mobile network [1], [2]. On the one hand, the use of THz communications is promising because a huge amount of bandwidth in the THz spectrum is available for communications [3]–[6]. On the other hand, the use of NOMA transmission can significantly improve spectral efficiency and support massive connectivity, by encouraging intelligent spectrum cooperation among mobile users [7], [8]. The two communication techniques are naturally complementary to each other. For example, using

Z. Ding and H. V. Poor are with the Department of Electrical and Computer Engineering, Princeton University, Princeton, NJ 08544, USA. Z. Ding is also with the School of Electrical and Electronic Engineering, the University of Manchester, Manchester, UK (email: zhiguo.ding@manchester.ac.uk, poor@princeton.edu).

NOMA to improve the spectral efficiency of THz networks is well motivated by the fact that some of the anticipated applications of 6G, such as immersive augmented reality (AR) and virtual reality (VR) as well as wireless transmission of ultra-high definition (UHD) video, can soon make the THz spectrum as crowded as those sub-6G Hz bands.

In the literature, THz-NOMA has been investigated from the following two perspectives. From the performance analysis perspective, the bit-error-rate performance of THz-NOMA has been studied in [9], where sophisticated schemes for adaptive superposition coding and subspace detection have been developed. In [10], NOMA has been applied to THz networks to mitigate beam misalignment errors, where the analytical results have been developed to show that the use of NOMA can improve the outage performance and user connectivity simultaneously. THz-NOMA can also be applied to cooperative communications in order to improve the coverage of mobile networks, as shown in [11]. From the resource allocation perspective, the issue of user clustering, i.e., which THz users are to be grouped together for the implementation of NOMA, has been studied in [12]–[14]. The application of advanced machine learning methods to resource allocation in THz-NOMA networks has also been investigated in [15], where intelligent reflecting surfaces have been used to reconfigure the wireless propagation environment.

Unlike these existing works about THz-NOMA, this paper focuses on how to use NOMA as a type of add-on in THz networks. In particular, in this paper, it is assumed that there exists a legacy THz network prior to the application of NOMA, where spatial beams have already been configured to serve legacy primary users. The aim of this paper is to investigate how to use these existing spatial beams and serve additional secondary users without degrading the performance of the legacy network. There are two motivations for the considered scenario. One motivation is that showing NOMA can be used as an add-on for existing communication systems demonstrates the compatibility of NOMA with other communication techniques and opens the door to more applications for NOMA [16]. The other motivation is to demonstrate that spatial beams can be used as a type of bandwidth resources, where allocating the pre-configured beams to secondary users in THz-NOMA networks is similar to subcarrier allocation in orthogonal frequency division multiplexing (OFDM) systems [17], [18]. The contributions of this paper are listed as follows:

- The considered joint beam management and power allocation problem is first formulated as a mixed combinatorial non-convex optimization problem, and then transformed to an equivalent but more concise matrix form, where the discrete optimization variables are eliminated and the number of continuous optimization variables is also reduced by using

the feature of the considered communication problem.

- Analytical results are developed in the paper to illustrate that the spatial beams can be used as bandwidth resources, but resource allocation based on these spatial beams is fundamentally different from methods based on conventional resources, such as OFDM subcarriers. In particular, because the pre-configured beams are not matched to the secondary users' channels, inter-beam interference exists, which makes conventional resource allocation approaches, such as water-filling power allocation, not applicable.
- The optimal performance of the considered THz-NOMA network is identified by applying the branch-and-bound (BB) method to the formulated resource allocation problem [18], [19]. Recall that the BB method can be viewed as a type of structured exhaustive search, which means that it yields the optimal performance but results in significant computational complexity. Therefore, low-complexity suboptimal resource allocation based on successive convex approximation (SCA) is also proposed in this paper [20], [21].
- Simulation results are also presented to illustrate some interesting features of beam based resource allocation. In particular, the use of NOMA can ensure that the overall system throughput of THz networks is significantly improved, by using those pre-configured spatial beams. For the special case with a single secondary user, greedy scheduling, i.e., using a single beam, is optimal. In addition, the throughput of the THz-NOMA network is improved if there are more secondary users involved because of multi-user diversity, but is degraded if the number of beams is increased because of inter-beam interference.

II. SYSTEM MODEL

For the considered THz communication system, there are two types of users, namely primary users and secondary users. The primary users form a legacy network, and the aim of the paper is to investigate how to serve those secondary users by using the spatial beams pre-configured for the primary users, as described in the following subsections.

A. A Legacy THz Network Based on Hybrid Beamforming

In the considered legacy network, K single-antenna primary users, denoted by U_k^P , $1 \leq k \leq K$, are served by a base station equipped with N antennas, where $K \leq N$ is assumed. Denote s_k^P by the signal to be sent to primary user U_k^P , and ρ_k^P by the corresponding transmit power.

Hybrid beamforming is used to serve these K primary users, i.e., the following signal vector is sent by the base station:

$$\tilde{\mathbf{s}}^P = \begin{bmatrix} \tilde{\mathbf{f}}_1 & \cdots & \tilde{\mathbf{f}}_K \end{bmatrix} \mathbf{P} \mathbf{s}^P, \quad (1)$$

where $\mathbf{s}^P = [s_1^P \ \cdots \ s_K^P]^T$, \mathbf{P} denotes the $K \times K$ digital beamforming matrix, and $\tilde{\mathbf{f}}_k$ denotes the analog beamforming vector for U_k^P .

There have been extensive studies for the design of hybrid beamforming. For example, beamsteering can be used to design analog beamforming [22], [23]. In particular, $\tilde{\mathbf{f}}_k$ can be selected from the following beamsteering codebook:

$$\tilde{\mathbf{f}}_k \in \left\{ \frac{1}{\sqrt{N}} \mathbf{a} \left(\frac{2\pi \times 0}{N_Q} \right), \dots, \frac{1}{\sqrt{N}} \mathbf{a} \left(\frac{2\pi(N_Q - 1)}{N_Q} \right) \right\}, \quad (2)$$

where N_Q denotes the size of the codebook, $\mathbf{a}(\theta)$ is an $N \times 1$ vector defined as follows:

$$\mathbf{a}(\theta) = \left[1 \quad e^{-j \frac{2\pi f_c d \sin(\theta)}{c}} \quad \cdots \quad e^{-j \frac{2(N-1)\pi f_c d \sin(\theta)}{c}} \right]^T, \quad (3)$$

f_c denotes the carrier frequency, d denotes the antenna spacing, and c denotes the speed of light. Therefore, $\tilde{\mathbf{f}}_k$ can be obtained by finding a vector from the codebook whose θ is closest to U_k^P 's angle of departure. After analog beamforming is obtained, simple approaches, such as zero forcing, can be used to design digital beamforming.

In this paper, it is assumed that both $\tilde{\mathbf{f}}_k$ and \mathbf{P} have already been configured prior to the implementation of NOMA, and the aim of the paper is to investigate how to serve additional users without changing the configuration of the legacy network, as discussed in the next subsection.

B. Serving Additional Users via THz-NOMA

Consider that there are M secondary users to be served via THz-NOMA. Denote the composite beamforming vector designed for primary user U_k^P by \mathbf{f}_k , i.e.,

$$\mathbf{f}_k = \begin{bmatrix} \tilde{\mathbf{f}}_1 & \cdots & \tilde{\mathbf{f}}_K \end{bmatrix} \mathbf{p}_k, \quad (4)$$

where \mathbf{p}_k denotes the k -th column of \mathbf{P} .

Each of the pre-configured K beams, \mathbf{f}_k , can be viewed as a type of bandwidth resources, and are to be allocated to the secondary users, which is similar to conventional subcarrier allocation problems in OFDMA systems. To facilitate the problem formulation, the beam allocation indicator, denoted by s_{jk} , is introduced [17], [24]. In particular, $s_{jk} = 1$ if secondary user U_j^S is allocated to the beam designed for primary user U_k^P , otherwise $s_{jk} = 0$. In order to reduce

the system complexity, it is assumed that at most one secondary user can be scheduled on each of the existing beams, \mathbf{f}_k , which imposes the following constraints on s_{jk} :

$$\sum_{j=1}^M s_{jk} = 1, \quad \& \quad s_{jk} \in \{0, 1\}, \quad \forall j, k. \quad (5)$$

By using the beam allocation indicator, U_k^P receives the following signal: [6], [22]

$$y_k^P = \frac{a_k^P}{\sqrt{\text{PL}_k^P}} \mathbf{a}^H(\theta_k^P) \sum_{i=1}^K \mathbf{f}_i \left(\sqrt{\rho_i^P} s_i^P + \sum_{m=1}^M s_{mi} \sqrt{\rho_{mi}^S} s_{mi}^S \right) + n_k^P, \quad (6)$$

where PL_k^P denotes the path loss suffered by U_k^P and is defined as follows:

$$\text{PL}_k^P = \left(\frac{c}{4\pi f_c} \right)^{-2} e^{\zeta r_{k,P}} \left(r_{k,P}^{\alpha_{\text{PL}}} + 1 \right), \quad (7)$$

$r_{k,P}$ denotes the distance between the base station and U_k^P , α_{PL} denotes the path loss exponent, ζ denotes the molecular absorption coefficient, a_k^P denotes the fading coefficient, θ_k^P denotes U_k^P 's angle of departure, s_{mi}^S denotes U_m^S 's signal sent on U_i^P 's beam, ρ_{mi}^S denotes U_m^S 's transmit power for signal s_{mi}^S , and n_k^P denotes the additive white Gaussian noise with power σ^2 .

To avoid changing the legacy system, it is assumed that the primary users treat the secondary users' signals as noise, and directly decode their own information, which means that the following data rate is achievable at the k -th primary user:

$$R_k^P = \log \left(1 + \frac{\frac{|a_k^P|^2}{\text{PL}_k^P} |\mathbf{a}^H(\theta_k^P) \mathbf{f}_k|^2 \rho_k^P}{\frac{|a_k^P|^2}{\text{PL}_k^P} |\mathbf{a}^H(\theta_k^P) \mathbf{f}_k|^2 \sum_{m=1}^M s_{mk} \rho_{mk}^S + I_{\text{IBI}}^k + \sigma^2} \right), \quad (8)$$

where I_{IBI} denotes inter-beam interference and is given by

$$I_{\text{IBI}}^k = \frac{|a_k^P|^2}{\text{PL}_k^P} \sum_{i=1, i \neq k}^K |\mathbf{a}^H(\theta_k^P) \mathbf{f}_i|^2 \left(\rho_i^P + \sum_{m=1}^M s_{mi} \rho_{mi}^S \right). \quad (9)$$

In order to guarantee U_k^P 's target data rate which is denoted by \bar{R}_k^P , the beam and power allocation for the secondary users should satisfy the following condition: $R_k^P \geq \bar{R}_k^P$.

On the other hand, if $s_{jk} = 1$, i.e., secondary user U_j^S is served on \mathbf{f}_k , this secondary user can decode the primary user's signal with the following data rate:

$$R_{j \rightarrow k}^S = \log \left(1 + \frac{\frac{|a_j^S|^2}{\text{PL}_j^S} |\mathbf{a}^H(\theta_j^S) \mathbf{f}_k|^2 \rho_k^P}{\frac{|a_j^S|^2}{\text{PL}_j^S} |\mathbf{a}^H(\theta_j^S) \mathbf{f}_k|^2 \sum_{m=1}^M s_{mk} \rho_{mk}^S + I_{\text{IBI}}^{jk} + \sigma^2} \right), \quad (10)$$

where the secondary user's channel parameters are defined similarly to those for the primary users and their definitions are omitted due to space limitations. The inter-beam interference, denoted by I_{IBI}^{jk} , is given by

$$I_{\text{IBI}}^{jk} = \frac{|a_j^S|^2}{\text{PL}_j^S} \sum_{i=1, i \neq k}^K |\mathbf{a}^H(\theta_j^S) \mathbf{f}_i|^2 \left(\rho_i^P + \sum_{m=1}^M s_{mi} \rho_{mi}^S \right). \quad (11)$$

Provided that $s_{jk} = 1$ and $R_{j \rightarrow k}^S \geq \bar{R}_k^P$, secondary user U_j^S can successfully decode primary user U_k^P 's signal and then decode its own signal sent on beam \mathbf{f}_k with the following data rate:

$$R_{j,k}^S = \log \left(1 + \frac{\frac{|a_j^S|^2}{\text{PL}_j^S} |\mathbf{a}^H(\theta_j^S) \mathbf{f}_k|^2 \rho_{jk}^S}{I_{\text{IBI}}^{jk} + \sigma^2} \right). \quad (12)$$

The aim of this paper is to design a joint beam management and power allocation approach for maximizing the secondary users' sum data rate, as formulated in the following:

$$\max_{\rho_{jk}^S, s_{jk}} \sum_{j=1}^M \sum_{k=1}^K s_{jk} R_{j,k}^S \quad (\text{P1a})$$

$$\text{s.t. } s_{jk} (R_{k,j}^P - \bar{R}_k^P) \geq 0, \forall k, j \quad (\text{P1b})$$

$$\tilde{R}_k^P - \bar{R}_k^P \geq 0, \forall k \quad (\text{P1c})$$

$$s_{jk} (R_{j \rightarrow k}^S - \bar{R}_k^P) \geq 0, \forall k, j \quad (\text{P1d})$$

$$s_{jk} \in \{0, 1\}, \forall k, j, \quad \sum_{j=1}^M s_{jk} \leq 1, \forall k \quad (\text{P1e})$$

$$\sum_{j=1}^M \sum_{k=1}^K s_{jk} \rho_{jk}^S \leq P^{\max}, \quad (\text{P1f})$$

where $\tilde{R}_k^P = \log \left(1 + \frac{\frac{|a_k^P|^2}{\text{PL}_k^P} |\mathbf{a}^H(\theta_k^P) \mathbf{f}_k|^2 \rho_k^P}{I_{\text{IBI}}^k + \sigma^2} \right)$, and P^{\max} denotes the transmit power budget for the secondary users. The constraints in (P1b) and (P1c) ensure that the primary users' quality of service (QoS) requirements, i.e., their target data rates, can be met when additional secondary users are served on the existing K beams. Note that in addition to constraint (P1b), constraint (P1c) is required because it is possible that no secondary user is scheduled on beam \mathbf{f}_k but primary user U_k^P still suffers the interference from the secondary users on the other beams. Constraint (P1c) can be omitted if the beams are orthogonal, i.e., $\mathbf{f}_k^H \mathbf{f}_i = 0, k \neq i$.

The constraint in (P1d) ensures that for a secondary user which is scheduled on beam \mathbf{f}_k , successive interference cancellation (SIC) can be carried out successfully. The constraint in

(P1e) ensures that at most one secondary user is served on each of the K beams. It is important to point out that it is possible that none of the secondary users is scheduled on one beam, and one secondary user is scheduled on multiple beams, i.e., the secondary users are scheduled in an opportunistic manner, where the investigation of user fairness among the secondary users is beyond the scope of this paper and will be treated as a promising direction for future research.

Problem P1 is challenging to solve since it is a mixed combinatorial non-convex optimization problem. In particular, it is straightforward to verify that the objective function is not concave, and the constraints in (P1b) and (P1d) are not convex. In addition, the beam allocation indicator, s_{jk} , is a binary optimization variable. In this paper, problem P1 will be solved by applying the BB and SCA methods which realize different performance-complexity tradeoffs.

III. PROBLEM REFORMULATION

In this section, the joint beam and power allocation problem presented in (P1) will be reformulated to facilitate the applications of the BB and SCA methods, where the property of the considered optimization problem is also studied.

In order to simplify the notations, first define $h_{ki}^P \triangleq \frac{|a_k^P|^2}{PL_k^P} |\mathbf{a}^H(\theta_k^P) \mathbf{f}_i|^2$, and R_k^P can be rewritten as follows:

$$R_k^P = \log \left(1 + \frac{h_{kk}^P \rho_k^P}{h_{kk}^P \sum_{m=1}^M s_{mk} \rho_{mk}^S + \sum_{i=1, i \neq k}^K h_{ki}^P (\rho_i^P + \sum_{m=1}^M s_{mi} \rho_{mi}^S) + \sigma^2} \right). \quad (13)$$

By using this simplified expression of R_k^P , constraint (P1b) can be simplified as follows:

$$s_{jk} \left(s_{jk} \rho_{jk}^S - \frac{\rho_k^P}{2^{\bar{R}_k^P} - 1} + \frac{1}{h_{kk}^P} \sum_{i=1, i \neq k}^K h_{ki}^P \left(\rho_i^P + \sum_{m=1}^M s_{mi} \rho_{mi}^S \right) + \frac{\sigma^2}{h_{kk}^P} \right) \leq 0, \quad (14)$$

where $\sum_{m=1}^M s_{mk} \rho_{mk}^S$ is reduced to $s_{jk} \rho_{jk}^S$ because of the use of s_{jk} outside of the bracket at the left-hand side of (14). In particular, if $s_{jk} = 1$, i.e., secondary user U_j^S is scheduled on beam \mathbf{f}_k , $\sum_{m=1}^M s_{mk} \rho_{mk}^S = s_{jk} \rho_{jk}^S$. If $s_{jk} = 0$, the constraint shown in (14) is not active, and the expression inside of the bracket at the left-hand side of (14) has no impact.

Note that ρ_k^P is a constant because the primary users' transmit powers are assumed to be fixed. Therefore, constraint (P1b) can be further simplified as follows:

$$s_{jk} \left(s_{jk} \rho_{jk}^S + \frac{1}{h_{kk}^P} \sum_{i=1, i \neq k}^K h_{ki}^P \sum_{m=1}^M s_{mi} \rho_{mi}^S + c_k \right) \leq 0, \quad (15)$$

where $c_k = \frac{1}{h_{kk}^P} \sum_{i=1, i \neq k}^K h_{ki}^P \rho_i^P - \frac{\rho_k^P}{2^{\bar{R}_k^P} - 1} + \frac{\sigma^2}{h_{kk}^P}$.

Similarly by introducing the following definition, $h_{jk}^S = \frac{|a_j^S|^2}{\text{PL}_j^S} |\mathbf{a}^H(\theta_j^S) \mathbf{f}_k|^2$, the data rate for U_j^S to decode U_k^P can be simplified as follows:

$$R_{j \rightarrow k}^S = \log \left(1 + \frac{h_{jk}^S \rho_k^P}{h_{jk}^S \sum_{m=1}^M s_{mk} \rho_{mk}^S + \sum_{i=1, i \neq k}^K h_{ji}^S (\rho_i^P + \sum_{m=1}^M s_{mi} \rho_{mi}^S) + \sigma^2} \right), \quad (16)$$

which means that constraint (P1d) which ensures the condition $R_{j \rightarrow k}^S \geq \bar{R}_k^P$ can be rewritten as follows:

$$s_{jk} \left(s_{jk} \rho_{jk}^S - \frac{\rho_k^P}{2^{\bar{R}_k^P} - 1} + \frac{1}{h_{jk}^S} \sum_{i=1, i \neq k}^K h_{ji}^S \left(\rho_i^P + \sum_{m=1}^M s_{mi} \rho_{mi}^S \right) + \frac{\sigma^2}{h_{jk}^S} \right) \leq 0, \quad (17)$$

where the beam allocation indicator is used to simplify the term $\sum_{m=1}^M s_{mk} \rho_{mk}^S$ to $s_{jk} \rho_{jk}^S$. By using the fact that the primary users' powers are constants, constraint (P1d) can be further simplified as follows:

$$s_{jk} \left(s_{jk} \rho_{jk}^S + \frac{1}{h_{jk}^S} \sum_{i=1, i \neq k}^K h_{ji}^S \sum_{m=1}^M s_{mi} \rho_{mi}^S + b_{jk} \right) \leq 0, \quad (18)$$

where $b_{jk} = \frac{1}{h_{jk}^S} \sum_{i=1, i \neq k}^K h_{ji}^S \rho_i^P - \frac{\rho_k^P}{2^{\bar{R}_k^P} - 1} + \frac{\sigma^2}{h_{jk}^S}$.

Similarly, by applying the above reformulation steps, the objective function can be also simplified as follows:

$$\begin{aligned} R_{j,k}^S &= \log \left(1 + \frac{h_{jk}^S \rho_{jk}^S}{\sum_{i=1, i \neq k}^K h_{ji}^S (\rho_i^P + \sum_{m=1}^M s_{mi} \rho_{mi}^S) + \sigma^2} \right) \\ &= \log \left(1 + \frac{h_{jk}^S \rho_{jk}^S}{\sum_{i=1, i \neq k}^K h_{ji}^S \sum_{m=1}^M s_{mi} \rho_{mi}^S + t_{jk}} \right), \end{aligned} \quad (19)$$

where $t_{jk} = \sum_{i=1, i \neq k}^K h_{ji}^S \rho_i^P + \sigma^2$.

By applying (15), (18) and (19) to problem P1, the considered optimization problem can be equivalently recast as follows:

$$\max_{\rho_{jk}^S, s_{jk}} \sum_{j=1}^M \sum_{k=1}^K s_{jk} \log \left(1 + \frac{h_{jk}^S \rho_{jk}^S}{\sum_{i=1, i \neq k}^K h_{ji}^S \sum_{m=1}^M s_{mi} \rho_{mi}^S + t_{jk}} \right) \quad (\text{P2a})$$

$$\text{s.t.} \quad s_{jk} \left(s_{jk} \rho_{jk}^S + \frac{1}{h_{kk}^P} \sum_{i=1, i \neq k}^K h_{ki}^P \sum_{m=1}^M s_{mi} \rho_{mi}^S + c_k \right) \leq 0, \forall k, j \quad (\text{P2b})$$

$$\frac{1}{h_{kk}^P} \sum_{i=1, i \neq k}^K h_{ki}^P \sum_{m=1}^M s_{mi} \rho_{mi}^S + c_k \leq 0, \forall k \quad (\text{P2c})$$

$$s_{jk} \left(s_{jk} \rho_{jk}^S + \frac{1}{h_{jk}^S} \sum_{i=1, i \neq k}^K h_{ji}^S \sum_{m=1}^M s_{mi} \rho_{mi}^S + b_{jk} \right) \leq 0, \forall k, j \quad (\text{P2d})$$

$$(\text{P1b}), (\text{P1e}), \& (\text{P1f}). \quad (\text{P2e})$$

Problem P2 is concise enough to obtain certain insight for the feature of beam-based resource allocation, as shown in the next subsection.

A. Special Cases with $M = 1$ and $K > 1$

When $M = 1$, there is a single secondary user and problem P2 is to find out how the overall transmit power, P^{\max} , can be distributed among the K beams. Intuitively, the water-filling approach can be applied, i.e., all the beams are employed and more power is allocated to a beam with a stronger channel gain. However, our conducted simulation results show a surprising result that greedy scheduling, i.e., using a single beam, is optimal. In particular, the greedy scheduling problem can be simply formulated as follows:

$$\max_{\rho_{1k}} \log \left(1 + \frac{h_{1k}^S}{t_{1k}} \max\{0, \min\{P^{\max}, -c_k, -b_{1k}\}\} \right). \quad (\text{P3a})$$

The following lemma shows that the conjecture that greedy scheduling is optimal holds in the special case with $M = 1$ and $K = 2$.

Lemma 1. *Consider a special case with $M = 1$ and $K = 2$, where the legacy network has been design to ensure that the target data rates of the primary users are small, i.e., $\bar{R}_k^P \rightarrow 0$, $k \in \{1, 2\}$, and all the primary users use the same transmit power. At high SNR, i.e., $\sigma^2 \rightarrow 0$, the optimal solution of problem P2 is the same as that of problem P3.*

Proof. See Appendix A. □

The proof for the conclusion that greedy scheduling is optimal for a more general case with $M = 1$ and $K > 1$ is difficult to obtain, and will be an important direction for future research. Note that for the general cases with $M > 1$, greedy scheduling is not optimal, and problem P2 needs to be further rearranged to facilitate the application of the BB and SCA methods, as shown in the next subsections.

B. Eliminating the Binary Optimization Variables, s_{jk}

Problem P2 is challenging to solve due to the facts that s_{jk} is binary and also the two optimization variables, s_{jk} and ρ_{jk}^S , are coupled. As shown in [24], the binary optimization variables can be eliminated by introducing the following continuous variable: $\tilde{\rho}_{jk}^S = s_{jk}\rho_{jk}^S$.

By using this auxiliary variable, constraints (P2b) and (P2c) can be combined together and equivalently expressed as follows:

$$\tilde{\rho}_{jk}^S + \frac{1}{h_{kk}^P} \sum_{i=1, i \neq k}^K h_{ki}^P \sum_{m=1}^M \tilde{\rho}_{mi}^S + c_k \leq 0, \forall k, j, \quad (20)$$

which can be explained in the following. For the case that $s_{jk} = 1$, $\tilde{\rho}_{jk}^S = \tilde{\rho}_{jk}^S$, and it is straightforward to show that (20) is equivalent to (P2b), which is stricter than (P2c) and hence constraint (P2c) can be ignored in this case. For the case that $s_{jk} = 0$, $\tilde{\rho}_{jk}^S = 0$, (20) is the same as (P2c), whereas constraint (P2b) is not active in this case.

Intuitively, constraints (P2b) can also be equivalently reformulated to the following concise expression by using $\tilde{\rho}_{jk}^S$:

$$\tilde{\rho}_{jk}^S + \frac{1}{h_{jk}^S} \sum_{i=1, i \neq k}^K h_{ji}^S \sum_{m=1}^M \tilde{\rho}_{mi}^S + b_{jk} \leq 0, \forall k, j. \quad (21)$$

For the case that $s_{jk} = 1$, (21) is indeed equivalent to (P2b). However, for the case that $s_{jk} = 0$, (21) is not equivalent to (P2b), because the original constraint in (P2b) is not active in this case but the new constraint in (21) is still active and expressed as follows: $\frac{1}{h_{jk}^S} \sum_{i=1, i \neq k}^K h_{ji}^S \sum_{m=1}^M \tilde{\rho}_{mi}^S + b_{jk} \leq 0$. Or in other words, the constraint in (21) cannot be used to replace (P2b) because an extra constraint is introduced if secondary user U_j^S is not scheduled on beam \mathbf{f}_k . Instead, constraint (P2b) can be equivalently recast as follows:

$$\text{sign}(\tilde{\rho}_{jk}^S) \left(\tilde{\rho}_{jk}^S + \frac{1}{h_{jk}^S} \sum_{i=1, i \neq k}^K h_{ji}^S \sum_{m=1}^M \tilde{\rho}_{mi}^S + b_{jk} \right) \leq 0, \forall k, j, \quad (22)$$

where $\text{sign}(x)$ denotes the sign of x .

Furthermore, $\tilde{\rho}_{jk}^S$ can also be used to simplify the objective function, where the following equality can be established:

$$s_{jk} \log \left(1 + \frac{h_{jk}^S \rho_{jk}^S}{\sum_{i=1, i \neq k}^K h_{ji}^S \sum_{m=1}^M s_{mi} \rho_{mi}^S + t_{jk}} \right) = \log \left(1 + \frac{h_{jk}^S \tilde{\rho}_{jk}^S}{\sum_{i=1, i \neq k}^K h_{ji}^S \sum_{m=1}^M \tilde{\rho}_{mi}^S + t_{jk}} \right), \quad (23)$$

which is explained in the following. For the case that $s_{jk} = 1$, $\tilde{\rho}_{jk}^S = \rho_{jk}^S$ and hence the two sides of (23) are the same. For the case that $s_{jk} = 0$, the two sides of (23) are zero and still equivalent.

Note that the use of these new expressions shown in (20) (22), and (23) can avoid using the beam assignment indicator. However, in order to ensure that at most a single secondary user

is scheduled on one beam, i.e., $\sum_{j=1}^M s_{jk} = 1$, a penalty variable, denoted by ξ , needs to be introduced, and the new objective function is given by

$$\sum_{j=1}^M \sum_{k=1}^K \log \left(1 + \frac{h_{jk}^S \tilde{\rho}_{jk}^S}{\xi h_{jk}^S \sum_{m=1, m \neq j}^M \tilde{\rho}_{mk}^S + \sum_{i=1, i \neq k}^K h_{ji}^S \sum_{m=1}^M \tilde{\rho}_{mi}^S + t_{jk}} \right). \quad (24)$$

By using (20) (22), and (24), problem P2 can be recast as follows:¹

$$\max_{\rho_{jk}^S} \sum_{j=1}^M \sum_{k=1}^K \log \left(1 + \frac{h_{jk}^S \tilde{\rho}_{jk}^S}{\xi h_{jk}^S \sum_{m=1, m \neq j}^M \tilde{\rho}_{mk}^S + \sum_{i=1, i \neq k}^K h_{ji}^S \sum_{m=1}^M \tilde{\rho}_{mi}^S + t_{jk}} \right) \quad (P4a)$$

$$\text{s.t.} \quad \sum_{j=1}^M \sum_{k=1}^K \tilde{\rho}_{jk}^S \leq P^{\max} \quad (P4b)$$

$$(20), (22). \quad (P4c)$$

As shown in [24], for the case $\xi \rightarrow \infty$, problem P4 is equivalent to problem P2.

C. Reducing the Number of Variables to Be Optimized

For problem P4, there are MK variables to be optimized, i.e., $\tilde{\rho}_{jk}$, $1 \leq j \leq M$ and $1 \leq k \leq K$, which can cause significant computational complexity. It is important to point out that the constraints in (20) and (22) can be used to reduce the number of optimization variables. For example, by using the constraint in (20), one can conclude that if $c_k > 0$, $\tilde{\rho}_{jk}^S = 0$ which also means $s_{jk} = 0$, $\forall j \in \{1, \dots, M\}$, because the following inequality can never be satisfied

$$\tilde{\rho}_{jk}^S + \frac{1}{h_{kk}^P} \sum_{i=1, i \neq k}^K h_{ki}^P \sum_{m=1}^M \tilde{\rho}_{mi}^S + c_k \leq 0. \quad (25)$$

This conclusion is also expected as explained in the following. Recall that $c_k = \frac{1}{h_{kk}^P} \sum_{i=1, i \neq k}^K h_{ki}^P \rho_i^P - \frac{\rho_k^P}{2^{\bar{R}_k^P} - 1} + \frac{\sigma^2}{h_{kk}^P}$, and hence $c_k > 0$ is equivalent to the following:

$$\log \left(1 + \frac{h_{kk}^P \rho_k^P}{\sum_{i=1, i \neq k}^K h_{ki}^P \rho_i^P + \sigma^2} \right) < \bar{R}_k^P, \quad (26)$$

which means that U_k^P 's QoS requirement cannot be satisfied even if no secondary user is served on beam \mathbf{f}_k . Or in other words, if $c_k > 0$, beam \mathbf{f}_k is not available to any secondary users. Similarly, by using the constraint in (22), one can conclude that $b_{jk} > 0$ leads to $\tilde{\rho}_{jk}^S = 0$ which also means $s_{jk} = 0$.

¹It can be straightforwardly verified that the objective and the constraints of problem P4 are monotonic functions, which means that similar to the BB method, monotonic optimization can also be used to find the optimal solution of problem P4.

In order to use the two conclusions for reducing the number of the optimization variables, first build the following set:

$$\mathcal{S} = \{\{j, k\} | b_{jk} \leq 0, c_k \leq 0\}. \quad (27)$$

Based on the previous discussions, $\tilde{\rho}_{jk} = 0$ and $s_{jk} = 0$, if $\{j, k\} \notin \mathcal{S}$, and hence there is no need to optimize these variables. Therefore, by using the set \mathcal{S} , problem P4 can be equivalently expressed as follows:

$$\max_{\rho_{jk}^S, \forall \{j, k\} \in \mathcal{S}} \sum_{\forall \{j, k\} \in \mathcal{S}} \log \left(1 + \frac{h_{jk}^S \tilde{\rho}_{jk}^S}{\xi h_{jk}^S \sum_{m=1, m \neq j}^M \tilde{\rho}_{mk}^S + \sum_{i=1, i \neq k}^K h_{ji}^S \sum_{m=1}^M \tilde{\rho}_{mi}^S + t_{jk}} \right) \quad (P5a)$$

$$\text{s.t.} \quad \tilde{\rho}_{jk}^S + \frac{1}{h_{kk}^P} \sum_{i=1, i \neq k}^K h_{ki}^P \sum_{m=1}^M \tilde{\rho}_{mi}^S + c_k \leq 0, \forall \{j, k\} \in \mathcal{S} \quad (P5b)$$

$$\text{sign}(\tilde{\rho}_{jk}^S) \left(\tilde{\rho}_{jk}^S + \frac{1}{h_{jk}^S} \sum_{i=1, i \neq k}^K h_{ji}^S \sum_{m=1}^M \tilde{\rho}_{mi}^S + b_{jk} \right) \leq 0, \forall \{j, k\} \in \mathcal{S} \quad (P5c)$$

$$\sum_{\forall \{j, k\} \in \mathcal{S}} \tilde{\rho}_{jk}^S \leq P^{\max}. \quad (P5d)$$

D. Reformulating the Problem into a Matrix-Based Form

In this subsection, problem P5 will be reformulated into a matrix form to facilitate the application of the BB and SCA methods. Recall that the set, \mathcal{S} , contains the indices of the secondary users whose transmit powers can be non-zero and need to be optimized, i.e., if $\{j, k\} \in \mathcal{S}$, secondary user U_j^S can be scheduled on beam \mathbf{f}_k . Denote the size of \mathcal{S} by $|\mathcal{S}|$. By using \mathcal{S} , one can build a $|\mathcal{S}| \times 2$ matrix, denoted by \mathbf{S} , where the two elements on the i -th row of \mathbf{S} are the i -th element of \mathcal{S} . For example, if $\mathcal{S} = \{\{1, 1\}, \{1, 2\}, \{2, 2\}\}$, $\mathbf{S} = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 2 \end{bmatrix}^T$.

Furthermore, denote the element on the p -th row and q -th column of a matrix \mathbf{S} by \mathbf{S}_{pq} . Define $y_p = \tilde{\rho}_{\mathbf{S}_{p1}\mathbf{S}_{p2}}$, and $\mathbf{y} = [y_1 \ \cdots \ y_{|\mathcal{S}|}]^T$ which collects all the variables to be optimized. Define $\tilde{\boldsymbol{\rho}}$ as a $KM \times 1$ vector collecting the original KM variables, $\tilde{\rho}_{jk}$, i.e.,

$$\tilde{\boldsymbol{\rho}} = [\tilde{\rho}_{11} \ \cdots \ \tilde{\rho}_{M1} \ \cdots \ \tilde{\rho}_{1K} \ \cdots \ \tilde{\rho}_{MK}]^T. \quad (28)$$

Furthermore, define \mathbf{R} as a $KM \times |\mathcal{S}|$ mapping matrix to ensure $\tilde{\boldsymbol{\rho}} = \mathbf{R}\mathbf{y}$, where \mathbf{R} can be built as follows. \mathbf{R} is an all zero matrix, except that the element on the $(M(\mathbf{S}_{p2} - 1) + \mathbf{S}_{p1})$ -th row and p -th column of \mathbf{R} , $1 \leq p \leq |\mathcal{S}|$, is set as one. For the above example with

$\mathcal{S} = \{\{1, 1\}, \{1, 2\}, \{2, 2\}\}$, $\mathbf{y} = [y_1 \ y_2 \ y_3]^T$, $y_1 = \tilde{\rho}_{\mathbf{S}_{11}\mathbf{S}_{12}} = \tilde{\rho}_{11}$, $y_2 = \tilde{\rho}_{\mathbf{S}_{21}\mathbf{S}_{22}} = \tilde{\rho}_{12}$, $y_3 = \tilde{\rho}_{\mathbf{S}_{31}\mathbf{S}_{32}} = \tilde{\rho}_{22}$, $\tilde{\boldsymbol{\rho}} = [\tilde{\rho}_{11} \ \tilde{\rho}_{21} \ \tilde{\rho}_{12} \ \tilde{\rho}_{22}]^T$, and \mathbf{R} is defined as follows:

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (29)$$

By using the above definitions, problem P5 can be equivalently recast as follows:

$$\max_{\mathbf{y}} \sum_{p=1}^{|\mathcal{S}|} \log \left(1 + \frac{\mathbf{c}_p^T \mathbf{y}}{\mathbf{d}_p^T \mathbf{R} \mathbf{y} + t_{\mathbf{S}_{p1}\mathbf{S}_{p2}}} \right) \quad (\text{P6a})$$

$$\text{s.t.} \quad \mathbf{e}_p^T \mathbf{R} \mathbf{y} + c_{\mathbf{S}_{p2}} \leq 0, 1 \leq p \leq |\mathcal{S}| \quad (\text{P6b})$$

$$\text{sign}(y_p) (\mathbf{f}_p^T \mathbf{R} \mathbf{y} + b_{\mathbf{S}_{p1}\mathbf{S}_{p2}}) \leq 0, 1 \leq p \leq |\mathcal{S}| \quad (\text{P6c})$$

$$\mathbf{1}_{|\mathcal{S}| \times 1}^T \mathbf{y} \leq P^{\max}. \quad (\text{P6d})$$

where \mathbf{c}_p is a $|\mathcal{S}| \times 1$ vector as follows:

$$\mathbf{c}_p = \left[\mathbf{0}_{1 \times (p-1)} \quad h_{\mathbf{S}_{p1}\mathbf{S}_{p2}}^S \quad \mathbf{0}_{1 \times (|\mathcal{S}|-p)} \right]^T, \quad (30)$$

\mathbf{d}_p is an $MK \times 1$ vector as follows:

$$\mathbf{d}_p = \left[h_{\mathbf{S}_{p11}}^S \mathbf{1}_{1 \times M} \quad \cdots \quad \xi h_{\mathbf{S}_{p1}\mathbf{S}_{p2}}^S \tilde{\mathbf{1}}_{1 \times M}^{\mathbf{S}_{p1}} \quad \cdots \quad h_{\mathbf{S}_{p1K}}^S \mathbf{1}_{1 \times M} \right]^T, \quad (31)$$

$\tilde{\mathbf{1}}_{1 \times M}^{\mathbf{S}_{p1}}$ is a $1 \times M$ all-one vector except that its \mathbf{S}_{p1} -th element is zero, \mathbf{e}_p is an $MK \times 1$ vector as follows:

$$\mathbf{e}_p = \left[\frac{h_{\mathbf{S}_{p21}}^P}{h_{\mathbf{S}_{p2}\mathbf{S}_{p2}}^P} \mathbf{1}_{1 \times M} \quad \cdots \quad \tilde{\mathbf{0}}_{1 \times M}^{\mathbf{S}_{p1}} \quad \cdots \quad \frac{h_{\mathbf{S}_{p2K}}^P}{h_{\mathbf{S}_{p2}\mathbf{S}_{p2}}^P} \mathbf{1}_{1 \times M} \right]^T, \quad (32)$$

$\tilde{\mathbf{0}}_{1 \times M}^{\mathbf{S}_{p1}}$ is a $1 \times M$ all-zero vector except that its \mathbf{S}_{p1} -th element is one, and \mathbf{f}_p is an $MK \times 1$ vector as follows:

$$\mathbf{f}_p = \left[\frac{h_{\mathbf{S}_{p11}}^S}{h_{\mathbf{S}_{p1}\mathbf{S}_{p2}}^S} \mathbf{1}_{1 \times M} \quad \cdots \quad \tilde{\mathbf{0}}_{1 \times M}^{\mathbf{S}_{p1}} \quad \cdots \quad \frac{h_{\mathbf{S}_{p1K}}^S}{h_{\mathbf{S}_{p1}\mathbf{S}_{p2}}^S} \mathbf{1}_{1 \times M} \right]^T. \quad (33)$$

The concise expression shown in (P6) provides insight to the considered joint beam and power allocation problem. For example, the constraints in (P6b) and (P6d) are based on simple affine functions. However, the constraint in (P6c) is not in a convex form due to the involvement of sign. In addition, the objective function is also not in a concave form as its logarithm terms contain ratios of linear functions.

IV. OPTIMAL AND SUBOPTIMAL SOLUTIONS FOR JOINT BEAM AND POWER ALLOCATION

In this section, two algorithms with different tradeoffs between performance and complexity are developed.

A. Applying the Branch and Bound Method

The BB method can be viewed as a type of structured search, where the feasibility region of the optimization problem is divided into smaller regions and a search for the optimal solution is carried out by focusing on those regions which are more promising and removing (pruning) those unlike ones. In the following, the considered optimization problem is first reformulated to facilitate the application of the BB method, and then the issue to find the upper and lower bounds on the optimal value of the considered optimization problem is focused.

1) *Implementation of the BB method:* Recall that the BB method can be ideally applied to the optimization problem with the following two features [18], [19]. One feature is that the feasibility region of the optimization problem can be expressed as a multi-dimensional rectangle, as it can be straightforwardly partitioned. The other feature is that the lower and upper bounds on the objective function for a partitioned feasibility region can be straightforwardly found. In order to recast the considered optimization problem to a form with the aforementioned two features, by introducing auxiliary variables, x_p , $1 \leq p \leq |\mathcal{S}|$, problem P6 can be first rewritten as follows:

$$\min_{\mathbf{y}, x_p} f(\mathbf{x}) = - \sum_{p=1}^{|\mathcal{S}|} \log(1 + x_p) \quad (\text{P7a})$$

$$\text{s.t. } x_p \leq \frac{\mathbf{c}_p^T \mathbf{y}}{\mathbf{d}_p^T \mathbf{R} \mathbf{y} + t_{\mathbf{s}_{p1}} \mathbf{s}_{p2}}, 1 \leq p \leq |\mathcal{S}| \quad (\text{P7b})$$

$$(\text{P6b}), (\text{P6c}), (\text{P6d}). \quad (\text{P7c})$$

Problem P7 facilitates the application of the BB method since its feasibility region with respect to x_p , i.e., (P7b), is a simple multi-dimensional rectangle. By defining $\mathbf{x} = [x_1 \ \cdots \ x_{|\mathcal{S}|}]$ and absorbing the variables, \mathbf{y} , into the feasibility region, problem P7 can be further rewritten as follows:

$$\min_{\mathbf{x}} f(\mathbf{x}) \quad \text{s.t. } \mathbf{x} \in \mathcal{G}. \quad (\text{P8a})$$

where

$$\mathcal{G} = \{\mathbf{x} | (\text{P7b}), (\text{P6b}), (\text{P6c}), (\text{P6d})\}. \quad (34)$$

Algorithm 1 Branch and Bound Algorithm

- 1: Set $\mathcal{B}_0 = \{\mathcal{D}_0\}$ and tolerance ϵ , $k = 0$, $U_0 = \phi^{\text{up}}(\mathcal{D}_0)$, $L_0 = \phi^{\text{lb}}(\mathcal{D}_0)$, and $\Delta = U_0 - L_0$
 - 2: **while** $\Delta \geq \epsilon$ **do**
 - 3: $k = k + 1$.
 - 4: Find $\mathcal{D} \in \mathcal{B}_{k-1}$ which yields the smallest lower bound
 - 5: Split \mathcal{D} along its longest edge into \mathcal{D}_1 and \mathcal{D}_2
 - 6: Construct $\mathcal{B}_k = \{\mathcal{D}_1 \cup \mathcal{D}_2 \cup (\mathcal{B}_{k-1} \setminus \mathcal{D})\}$
 - 7: Find the updated upper bound $U_k = \min \phi^{\text{up}}(\mathcal{D}), \forall \mathcal{D} \in \mathcal{B}_k$.
 - 8: Find the updated lower bound $L_k = \min \phi^{\text{lb}}(\mathcal{D}), \forall \mathcal{D} \in \mathcal{B}_k$.
 - 9: Update Δ by using $\Delta = U_k - L_k$
 - 10: Remove those \mathcal{D} in \mathcal{B}_k whose lower bounds are larger than U_k .
 - 11: **end**
-

As to be shown in the next subsection, the task to find the lower and upper on the optimal value can be simplified if the objective function of the considered optimization problem is a monotonically decreasing function, which motivates the objective function to be rewritten as follows:

$$\tilde{f}(\mathbf{x}) = \begin{cases} -\sum_{p=1}^{|\mathcal{S}|} \log(1 + x_p), & \text{if } \mathbf{x} \in \mathcal{G} \\ 0, & \text{otherwise} \end{cases}. \quad (35)$$

It is straightforward to verify that $\tilde{f}(\mathbf{x})$ is a monotonically decreasing function of \mathbf{x} , and problem P8 can be expressed as follows:

$$\min_{\mathbf{x}} \tilde{f}(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{x} \in \mathcal{G}. \quad (\text{P9a})$$

As shown in [18], problem P9 yields the same optimal solution as problem P8. Problem P9 can be solved efficiently by applying the BB method, as shown in Algorithm 1.

As shown in Algorithm 1, the initialization of the BB method needs to define an initial search region, which is a $|\mathcal{S}|$ -dimensional rectangle, denoted by \mathcal{D}_0 and defined as follows:

$$\mathcal{D}_0 = \{\mathbf{x} | 0 \leq x_p \leq x_p^B\}, \quad (36)$$

where $x_p^B = \frac{P_{\max} \mathbf{c}_p^T \mathbf{1}_{|\mathcal{S}| \times 1}}{t_{\mathbf{s}_{p1}} \mathbf{s}_{p2}}$ denotes the upper bound on x_p .

During the k -th iteration, the rectangle which yields the smallest lower bound among all the rectangles in the set, \mathcal{B}_{k-1} , is located and partitioned into two smaller rectangles along

its longest edge, denoted by \mathcal{D}_1 and \mathcal{D}_2 , respectively. For each rectangle, \mathcal{D}_i , find the upper and lower bounds on the optimal value of the considered optimization problem, denoted by $\phi^{\text{up}}(\mathcal{D}_i)$ and $\phi^{\text{lb}}(\mathcal{D}_i)$, respectively. More discussions will be provided in the next subsection for calculating $\phi^{\text{up}}(\mathcal{D})$ and $\phi^{\text{lb}}(\mathcal{D})$. The overall lower and upper bounds on the optimal value can be updated iteratively as shown in Algorithm 1, where this updated upper bound can also be used to remove those rectangles whose lower bounds are larger than the new upper bound. The algorithm terminates when the difference between the overall lower and upper bounds is smaller than the given tolerance parameter, denoted by ϵ .

2) *Finding the upper and lower bounds $\phi^{\text{up}}(\mathcal{D})$ and $\phi^{\text{lb}}(\mathcal{D})$* : For each rectangle, \mathcal{D} , denote its maximum and minimum vertices by \mathbf{x}_{max} and \mathbf{x}_{min} , respectively, i.e., $\mathbf{x}_{\text{min}} \leq \mathbf{x} \leq \mathbf{x}_{\text{max}}$ for $\mathbf{x} \in \mathcal{D}$. By using the fact that $\tilde{f}(\mathbf{x})$ is a monotonically decreasing function of \mathbf{x} , the lower and upper bounds, $\phi^{\text{lb}}(\mathcal{D})$ and $\phi^{\text{up}}(\mathcal{D})$, can be calculated as follows:

$$\phi^{\text{lb}}(\mathcal{D}) = \begin{cases} f(\mathbf{x}_{\text{max}}), & \text{if } \mathbf{x}_{\text{min}} \in \mathcal{G} \\ 0, & \text{otherwise} \end{cases}, \quad (37)$$

and

$$\phi^{\text{up}}(\mathcal{D}) = \begin{cases} f(\mathbf{x}_{\text{min}}), & \text{if } \mathbf{x}_{\text{min}} \in \mathcal{G} \\ 0, & \text{otherwise} \end{cases}. \quad (38)$$

The rationale behind the bounds shown in (37) and (38) is that, if \mathbf{x}_{min} is feasible, \mathbf{x}_{min} yields the maximal value for the objective function since it is the minimum element in \mathcal{D} , whereas \mathbf{x}_{max} is the maximum element in \mathcal{D} and can be used to find a (not necessarily achievable) lower bound on the objective function. If \mathbf{x}_{min} is not feasible, the rectangle is located outside of the feasibility region and hence should be pruned (removed), where the use of zero for the upper and lower bounds in this case can realize this goal. The task to verify whether $\mathbf{x}_{\text{min}} \in \mathcal{G}$, i.e., \mathbf{x}_{min} is feasible, can be accomplished by carrying out the following feasibility study

$$\max_{\mathbf{y}} \quad 1 \quad (\text{P10a})$$

$$\text{s.t.} \quad (\mathbf{c}_p^T - x_p \mathbf{d}_p^T \mathbf{R}) \mathbf{y} \geq x_{\text{min},p} t_{\mathbf{s}_{p1} \mathbf{s}_{p2}} \quad (\text{P10b})$$

$$\mathbf{e}_p^T \mathbf{R} \mathbf{y} + c_{\mathbf{s}_{p2}} \leq 0, 1 \leq p \leq |\mathcal{S}| \quad (\text{P10c})$$

$$\text{sign}(y_p) (\mathbf{f}_p^T \mathbf{R} \mathbf{y} + b_{\mathbf{s}_{p1} \mathbf{s}_{p2}}) \leq 0, 1 \leq p \leq |\mathcal{S}| \quad (\text{P10d})$$

$$\mathbf{1}_{|\mathcal{S}| \times 1}^T \mathbf{y} \leq P^{\text{max}}, \quad (\text{P10e})$$

where $x_{\min,p}$ is the p -th element of \mathbf{x}_{\min} . Note that the use of $\text{sign}(y_p)$ makes constraint (P10d) not convex. It is important to point out that $y_p = 0$ and $x_{\min,p} = 0$ are equivalent, i.e., $y_p = 0$ leads to $x_{\min,p} = 0$ and vice versa. By using this observation, problem P10 can be recast as the following equivalent form:

$$\max_{\mathbf{y}} \quad 1 \tag{P11a}$$

$$\text{s.t.} \quad \text{sign}(x_{\min,p}) (\mathbf{f}_p^T \mathbf{R} \mathbf{y} + b_{\mathbf{s}_{p1} \mathbf{s}_{p2}}) \leq 0, 1 \leq p \leq |\mathcal{S}| \tag{P11b}$$

$$\text{(P10b), (P10c), (P10e)}. \tag{P11c}$$

Note that in problem P11, \mathbf{x}_{\min} is not an optimization variable and hence $\text{sign}(x_{\min,p})$ is a constant, which means that constraint (P11b) is an affine function. Since all the constraint functions of problem P11 are affine, problem P11 can be solved efficiently by applying those off-shelf optimization solvers.

Because the convergency of the BB method depends on how tight the upper and lower bounds are. Therefore, the bounds shown in (37) and (38) will be further tightened, by using the steps shown in [19].

3) *Tightening the upper and lower bounds:* In this section, we will focus on the case that $\mathbf{x}_{\min} \in \mathcal{G}$, otherwise the corresponding rectangle will be eventually pruned. Furthermore, we assume that $\mathbf{x}_{\max} \notin \mathcal{G}$, otherwise the bounds shown in (37) and (38) are tight. In (37), a lower bound is obtained by directly using \mathbf{x}_{\max} which is often far away from the boundary of the feasibility region. The key idea for getting a lower bound tighter than (37) is to find a new vector, denoted by $\tilde{\mathbf{x}}_{\max}$, which is closer to the boundary of the feasibility region than \mathbf{x}_{\max} . In particular, the p -th element of $\tilde{\mathbf{x}}_{\max}$, denoted by $\tilde{x}_{\max,p}$, is obtained by finding its maximal value if all the other elements of $\tilde{\mathbf{x}}_{\max}$ are the same as those of \mathbf{x}_{\min} , which is to solve the following optimization problem

$$\max_{\mathbf{y}} \quad \tilde{x}_{\max,p} \triangleq \frac{\mathbf{c}_p^T \mathbf{y}}{\mathbf{d}_p^T \mathbf{R} \mathbf{y} + t_{\mathbf{s}_{p1} \mathbf{s}_{p2}}} \tag{P12a}$$

$$\text{s.t.} \quad \frac{\mathbf{c}_p^T \mathbf{y}}{\mathbf{d}_p^T \mathbf{R} \mathbf{y} + t_{\mathbf{s}_{p1} \mathbf{s}_{p2}}} \leq x_{\max,p} \tag{P12b}$$

$$x_{\min,i} = \frac{\mathbf{c}_i^T \mathbf{y}}{\mathbf{d}_i^T \mathbf{R} \mathbf{y} + t_{\mathbf{s}_{i1} \mathbf{s}_{i2}}}, i \neq p, 1 \leq i \leq |\mathcal{S}| \tag{P12c}$$

$$\text{(P10c), (P11b), (P10e)}. \tag{P12d}$$

Problem P12 can be rewritten as follows:

$$\max_{\mathbf{y}} \frac{\mathbf{c}_p^T \mathbf{y}}{\mathbf{d}_p^T \mathbf{R} \mathbf{y} + t_{\mathbf{s}_{p1} \mathbf{s}_{p2}}} \quad (\text{P13a})$$

$$\text{s.t.} \quad (\mathbf{c}_p^T - x_{\max,p} \mathbf{d}_p^T \mathbf{R}) \mathbf{y} \leq x_{\max,p} t_{\mathbf{s}_{p1} \mathbf{s}_{p2}} \quad (\text{P13b})$$

$$(\mathbf{c}_i^T - x_{\min,i} \mathbf{d}_i^T \mathbf{R}) \mathbf{y} = x_{\min,i} t_{\mathbf{s}_{i1} \mathbf{s}_{i2}}, i \neq p, 1 \leq i \leq |\mathcal{S}| \quad (\text{P13c})$$

$$\mathbf{e}_p^T \mathbf{R} \mathbf{y} \leq -c_{\mathbf{s}_{p2}}, 1 \leq p \leq |\mathcal{S}| \quad (\text{P13d})$$

$$x_{\min,p} \mathbf{f}_p^T \mathbf{R} \mathbf{y} \leq -x_{\min,p} b_{\mathbf{s}_{p1} \mathbf{s}_{p2}}, 1 \leq p \leq |\mathcal{S}| \quad (\text{P13e})$$

$$\mathbf{1}_{|\mathcal{S}| \times 1}^T \mathbf{y} \leq P^{\max}. \quad (\text{P13f})$$

Note that all the constraints of problem P13 are affine, and hence they can be grouped to yield the following more concise form:

$$\max_{\mathbf{y}} \frac{\mathbf{c}_p^T \mathbf{y}}{\mathbf{d}_p^T \mathbf{R} \mathbf{y} + t_{\mathbf{s}_{p1} \mathbf{s}_{p2}}} \quad (\text{P14a})$$

$$\text{s.t.} \quad \mathbf{A}_p \mathbf{y} \leq \mathbf{b}_p \quad (\text{P14b})$$

$$\mathbf{A}_p^E \mathbf{y} = \mathbf{b}_p^E, \quad (\text{P14c})$$

where the expressions for \mathbf{A}_p , \mathbf{b}_p , \mathbf{A}_p^E and \mathbf{b}_p^E can be straightforwardly obtained from problem P13 and are omitted due to space limitations.

Define $\tilde{\mathbf{A}}_p^E$ as a $(|\mathcal{S}| - 1) \times (|\mathcal{S}| - 1)$ square matrix obtained from \mathbf{A}_p^E by removing its p -th column of \mathbf{A}_p^E , denoted by $\mathbf{a}_{p,p}^E$. $\tilde{\mathbf{y}}_p$ is a $(|\mathcal{S}| - 1) \times 1$ vector obtained from \mathbf{y} by removing y_p . A close-form expression for the optimal solution of problem P14 can be obtained as follows.

Lemma 2. *Assume that $\tilde{\mathbf{A}}_p^E$ is invertible and problem P14 is feasible. The optimal solution for problem P14, denoted by \mathbf{y}_p^* , can be obtained as follows. First, the p -th element of \mathbf{y}_p^* can be expressed as follows:*

$$y_p^* = \min\{\mathbf{a}_{\text{sign}} \odot (\mathbf{b}_p - \tilde{\mathbf{A}}_p(\tilde{\mathbf{A}}_p^E)^{-1} \mathbf{b}_p^E) ./ (\mathbf{a}_{p,p} - \tilde{\mathbf{A}}_p(\tilde{\mathbf{A}}_p^E)^{-1} \mathbf{a}_{p,p}^E)\}, \quad (39)$$

where $\mathbf{a}_{p,p}$ denotes the p -th column of \mathbf{A}_p , $\tilde{\mathbf{A}}_p$ is obtained from \mathbf{A}_p by removing $\mathbf{a}_{p,p}$, $\mathbf{a}_{\text{sign}} = \max\{0, (\mathbf{a}_{p,p} - \tilde{\mathbf{A}}_p(\tilde{\mathbf{A}}_p^E)^{-1} \mathbf{a}_{p,p}^E)\}$, \odot and $./$ denote element-wise multiplication and division, respectively. Second, collect the remaining $(|\mathcal{S}| - 1)$ elements of \mathbf{y}_p^* in the vector, denoted by $\tilde{\mathbf{y}}_p^*$, and $\tilde{\mathbf{y}}_p^*$ can be obtained from y_p^* as follows: $\tilde{\mathbf{y}}_p^* = (\tilde{\mathbf{A}}_p^E)^{-1}(\mathbf{b}_p^E - \mathbf{a}_{p,p}^E y_p^*)$.

Proof. See Appendix B □

Among our conducted computer simulations, we notice that $\tilde{\mathbf{A}}_p^E$ can be close to singular for a small number of channel realizations. For these rare cases, problem P14 can still be solved efficiently by first reformulating it as a linear programming problem and then applying those optimization solvers. In particular, define $\mathbf{z} = \frac{\mathbf{y}}{\mathbf{d}_p^T \mathbf{R} \mathbf{y} + t_{\mathbf{s}_{p1}} \mathbf{s}_{p2}}$ and $w = \frac{1}{\mathbf{d}_p^T \mathbf{R} \mathbf{y} + t_{\mathbf{s}_{p1}} \mathbf{s}_{p2}}$, which means that problem P14 can be recast as follows:

$$\max_{\mathbf{y}} \quad \mathbf{c}_p^T \mathbf{z} \quad (\text{P15a})$$

$$\text{s.t.} \quad \mathbf{A}_p \mathbf{z} \leq w \mathbf{b}_p \quad (\text{P15b})$$

$$\mathbf{A}_p^E \mathbf{z} = w \mathbf{b}_p^E \quad (\text{P15c})$$

$$\mathbf{d}_p^T \mathbf{R} \mathbf{z} + t_{\mathbf{s}_{p1}} \mathbf{s}_{p2} w = 1. \quad (\text{P15d})$$

Problem P15 is in a linear programming form and hence can be directly solved by applying those off-shelf optimization solvers. Once all the elements of $\tilde{\mathbf{x}}_{\max}$ are obtained, a tighter lower bound can be found by replacing \mathbf{x}_{\max} with $\tilde{\mathbf{x}}_{\max}$ in (37).

Interestingly, $\tilde{\mathbf{x}}_{\max}$ can also be used to tighten the upper bound. Recall that the p -th element of $\tilde{\mathbf{x}}_{\max}$ is obtained by first assuming that the other elements of $\tilde{\mathbf{x}}_{\max}$ are equal to those in \mathbf{x}_{\min} and then solving problem P14. Therefore, build the $|\mathcal{S}|$ vectors, denoted by $\tilde{\mathbf{x}}_{\min}^i$, $1 \leq i \leq |\mathcal{S}|$, where each of the vectors is a $|\mathcal{S}| \times 1$ vector, and its p -th element, denoted by $\tilde{x}_{\min,p}^i$, is given by

$$\begin{cases} \tilde{x}_{\min,p}^i = \tilde{x}_{\max,p} & \text{if } p = i \\ \tilde{x}_{\min,p}^i = x_{\min,p} & \text{if } p \neq i \end{cases}. \quad (40)$$

According to the steps to find $\tilde{x}_{\max,p}$, all the vectors, $\tilde{\mathbf{x}}_{\min}^i$, $1 \leq i \leq |\mathcal{S}|$, are feasible, i.e., $\tilde{\mathbf{x}}_{\min}^i \in \mathcal{G}$. Therefore, these vectors can be used to form new upper bounds on the optimal value. In particular, by replacing \mathbf{x}_{\min}^i with $\tilde{\mathbf{x}}_{\min}^i$, $1 \leq i \leq |\mathcal{S}|$, in (38), $|\mathcal{S}|$ new upper bounds can be obtained, where the smallest one can be used as the tightened upper bound.

B. Applying the Successive Convex Approximation Method

Recall that the BB method is essentially a structured search, where many iterations are required in order to divide the feasibility region into sufficiently small multi-dimensional rectangles. As a result, the computational complexity of the BB method can be significant, particularly in the case that the number of optimization variables, $|\mathcal{S}|$, is large, which motivates the use of the SCA method.

In order to facilitate the application of SCA, problem P6 can be first re-written as follows:

$$\max_{\mathbf{y}} \sum_{p=1}^{|\mathcal{S}|} [\log(\mathbf{c}_p^T \mathbf{y} + \mathbf{d}_p^T \mathbf{R} \mathbf{y} + t_{\mathbf{s}_{p1} \mathbf{s}_{p2}}) - \log(\mathbf{d}_p^T \mathbf{R} \mathbf{y} + t_{\mathbf{s}_{p1} \mathbf{s}_{p2}})] \quad (\text{P16a})$$

$$\text{s.t. } \mathbf{e}_p^T \mathbf{R} \mathbf{y} + c_{\mathbf{s}_{p2}} \leq 0, 1 \leq p \leq |\mathcal{S}| \quad (\text{P16b})$$

$$\text{sign}(y_p) (\mathbf{f}_p^T \mathbf{R} \mathbf{y} + b_{\mathbf{s}_{p1} \mathbf{s}_{p2}}) \leq 0, 1 \leq p \leq |\mathcal{S}| \quad (\text{P16c})$$

$$\mathbf{1}_{|\mathcal{S}| \times 1}^T \mathbf{y} \leq P^{\max}. \quad (\text{P16d})$$

To tackle the challenge that the objective function of problem P16 is not concave, auxiliary optimization variables, z_p , are introduced and problem P16 can be equivalently recast as follows:

$$\max_{\mathbf{y}, z_p} \sum_{p=1}^{|\mathcal{S}|} \log(\mathbf{c}_p^T \mathbf{y} + \mathbf{d}_p^T \mathbf{R} \mathbf{y} + t_{\mathbf{s}_{p1} \mathbf{s}_{p2}}) - \sum_{p=1}^{|\mathcal{S}|} z_p \quad (\text{P17a})$$

$$\text{s.t. } \log(\mathbf{d}_p^T \mathbf{R} \mathbf{y} + t_{\mathbf{s}_{p1} \mathbf{s}_{p2}}) \leq z_p, 1 \leq p \leq |\mathcal{S}| \quad (\text{P17b})$$

$$(\text{P16b}), (\text{P16c}), (\text{P16d}). \quad (\text{P17c})$$

Note that the constraint in (P17b) is not convex, but it can be approximated by using the first order Taylor expansion, which means that problem P16 can be approximated as follows:

$$\max_{\mathbf{y}, z_p} \sum_{p=1}^{|\mathcal{S}|} \log(\mathbf{c}_p^T \mathbf{y} + \mathbf{d}_p^T \mathbf{R} \mathbf{y} + t_{\mathbf{s}_{p1} \mathbf{s}_{p2}}) - \sum_{p=1}^{|\mathcal{S}|} z_p \quad (\text{P18a})$$

$$\text{s.t. } \log(\mathbf{d}_p^T \mathbf{R} \mathbf{y}_0 + t_{\mathbf{s}_{p1} \mathbf{s}_{p2}}) + \frac{\mathbf{d}_p^T \mathbf{R} (\mathbf{y} - \mathbf{y}_0)}{\ln(2) (\mathbf{d}_p^T \mathbf{R} \mathbf{y}_0 + t_{\mathbf{s}_{p1} \mathbf{s}_{p2}})} \leq z_p, \forall p \quad (\text{P18b})$$

$$(\text{P16b}), (\text{P16c}), (\text{P16d}). \quad (\text{P18c})$$

where \mathbf{y}_0 denotes an initial estimate of \mathbf{y} and can be iteratively updated. It is straightforward to verify that the objective function of problem P18 is concave, and the newly introduced constraint in (P18b) is a simple affine function.

The only remaining challenge to solve problem P18 is that constraint (P16c) is still not in a convex form due to the use of the sign function. In the following, two heuristic solutions, termed SCA-I and SCA-II, respectively, are proposed to reformulate problem P18 into a concave

optimization form. SCA-I is to directly remove the sign function in (P16c), which leads to the following optimization problem:

$$\max_{\mathbf{y}, z_p} \sum_{p=1}^{|\mathcal{S}|} \log(\mathbf{c}_p^T \mathbf{y} + \mathbf{d}_p^T \mathbf{R} \mathbf{y} + t_{\mathbf{s}_{p1} \mathbf{s}_{p2}}) - \sum_{p=1}^{|\mathcal{S}|} z_p \quad (\text{P19a})$$

$$\text{s.t.} \quad (\mathbf{f}_p^T \mathbf{R} \mathbf{y} + b_{\mathbf{s}_{p1} \mathbf{s}_{p2}}) \leq 0, 1 \leq p \leq |\mathcal{S}| \quad (\text{P19b})$$

$$(\text{P16b}), (\text{P16d}), (\text{P18b}), \quad (\text{P19c})$$

which is a typical concave maximization problem, and can be solved efficiently by applying the optimization solvers.

SCA-II is motivated by the fact that the challenge in constraint (P16c) is caused by the use of the beam assignment indicator function. If beam assignment is carried out before power allocation, this challenging issue can be avoided. Therefore, SCA-II consists of two steps. The first step is to carry out user scheduling on each beam, i.e., secondary user $U_{j_k^*}^k$ is scheduled on beam \mathbf{f}_k if $j_k^* = \arg \max_j \{h_{j,k}^S, \{j, k\} \in \mathcal{S}\}$. The second step of SCA-II is to update \mathcal{S} by including $\{j_k^*, k\}$, $1 \leq k \leq K$, only, and then carry out power allocation, i.e., solving problem P19 in the same manner as SCA-I.

V. SIMULATION RESULTS

In this section, the computer simulation results are presented to evaluate the performance of THz-NOMA with joint beam and power allocation. Motivated by Lemma 1, the greedy scheduling scheme is used as a benchmarking scheme for the BB and SCA methods. For all conducted simulations, $\rho_P = 30$ dBm, $\sigma^2 = -90$ dBm, $P_{\max} = 30$ dBm, $\xi = 10^8$, $\alpha_{\text{PL}} = 2$, $f_c = 300$ GHz, $\zeta = 5e^{-3}$, $d = \frac{c}{2f_c}$, and $\phi = 0.1$, as in [6]. The primary users are uniformly deployed within a square with its edge length 10 m, where $\theta_k^P = \frac{i}{K}\pi - \frac{\pi}{2}$, $1 \leq i \leq K$. The secondary users are also uniformly deployed within a square with its edge length denoted by r_S , where θ_j^S is uniformly distributed between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$. Recall that the BB method can be viewed as a structured exhaustive search, and its convergence requires a significant number of iterations, particularly if there are a large number of users. A useful observation to reduce the complexity of the implementation of the BB method is that when sufficient iterations are carried out, the rectangles in \mathcal{B}_k are already small enough to provide a good estimate for the optimal value. Table I shows the effect of capping the number of iterations for the BB method, where N_{itr} denotes the maximal number of iterations, $N = 10$, $K = 4$, and $N_Q = 10$. As can be seen

TABLE I
IMPACT OF N_{itr} ON THE PERFORMANCE OF THE BB METHOD

	$M = 1$	$M = 2$	$M = 4$	$M = 6$	$M = 8$
$N_{itr} = \infty$	2.2805	4.04997	5.7922	6.9129	7.8640
$N_{itr} = 200$	2.2791	3.8855	5.7205	6.8343	7.4128

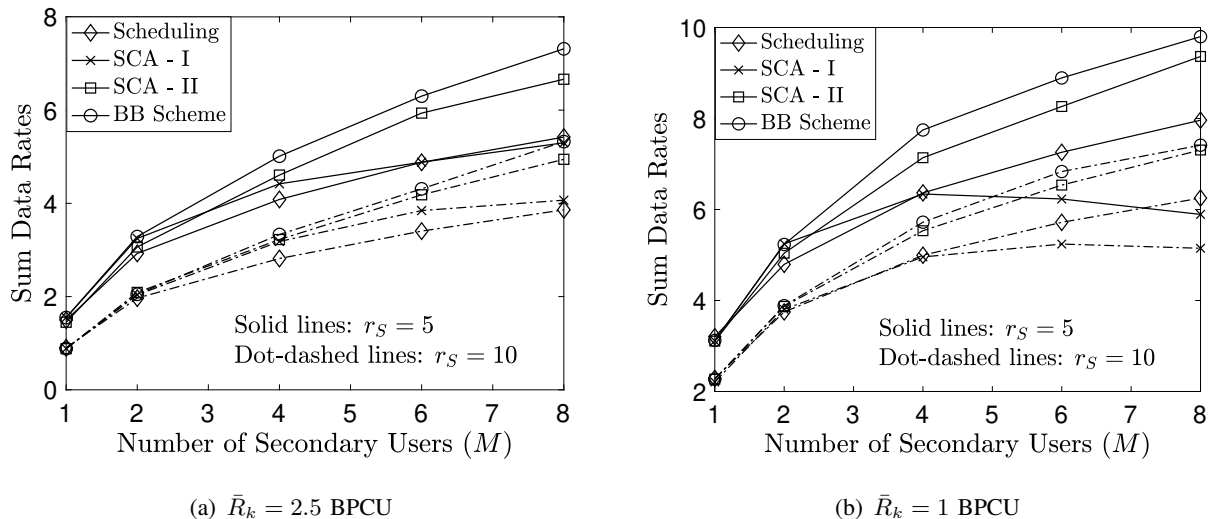


Fig. 1. Impact of the number of secondary users on the performance of THz-NOMA. $\alpha = 2$, $N = 10$, $K = 4$, $N_Q = 10$. Hybrid beamforming is used.

from the table, capping the number of iterations does not cause a significant performance loss for the BB method, and hence $N_{itr} = 200$ is used in the following conducted simulations.

In Fig. 1, the impact of the number of secondary users on the performance of THz-NOMA is studied, where different choices of r_S are used. As can be seen from Fig. 1, the use of THz-NOMA can ensure that the secondary users are served on those existing beams with significant data rates. This means that the overall system throughput of THz networks can be significantly improved compared to the case in which only the primary users are served. Fig. 1 also shows that the sum rate gain can be further increased by increasing M , or reducing r_S and \bar{R}_k . Among the considered schemes, the BB method yields the best performance, since it is a structured search and is expected to provide the optimal performance. When there is a single secondary user, the greedy scheduling scheme realizes the same performance as the BB method, which confirms Lemma 1. SCA-I is a naive application of the SCA method, and its performance can be even worse than the greedy scheduling scheme, particularly for the case with a larger M . SCA-II is the combination of user scheduling and SCA, and Fig. 1 shows that the SCA-II can outperform the greedy scheduling scheme, and realize a performance close to the optimal BB

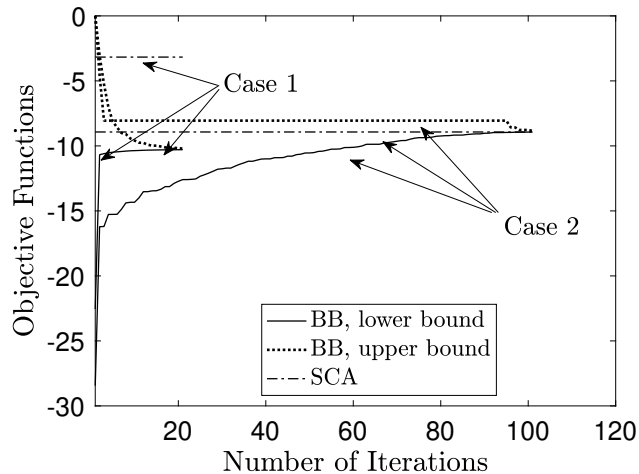


Fig. 2. Illustration of the convergence of the SCA-II and BB schemes. Two random realizations of channels are used. $N = 10$, $K = 4$, $M = 8$, $R_k = 2.5$ BPCU, $N_Q = 10$ and $r_S = 5$.

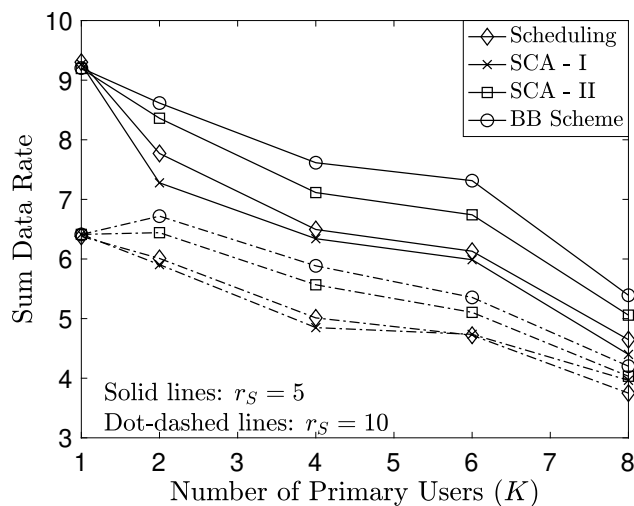


Fig. 3. Impact of the number of primary users on the performance of THz-NOMA transmission. $N = 10$, $M = 4$, $R_k = 1$ BPCU, $N_Q = 10$ and $r_S = 10$.

method. It is important to point out that the convergence of SCA is much faster than the BB method, as shown in Fig. 2, which means that the complexity of SCA is much smaller than that of the BB method. In particular, Fig. 2 demonstrates that SCA can converge within a single iteration, whereas the BB method can take hundreds of iterations to converge, even for the case with a moderate number of users.

In Fig. 3, the impact of the number of primary users on the performance of THz-NOMA transmission is studied. Recall that Fig. 1 shows that inviting more secondary users to participate in THz-NOMA transmission can increase the overall sum rate, because a larger secondary user pool is helpful to improve the effective channel gains of the scheduled secondary users. Intuitively, increasing K should also be helpful to increase the sum rate, since there are more beams, i.e., there are more bandwidth resources available. Fig. 3 shows a surprising result that

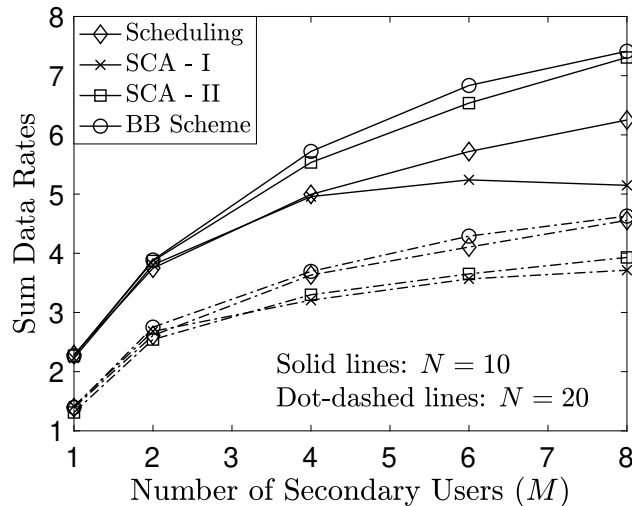


Fig. 4. Impact of the number of antennas at the base station on the performance of THz-NOMA transmission. $K = 4$, $R_k = 1$ BPCU, $N_Q = 10$ and $r_S = 10$.

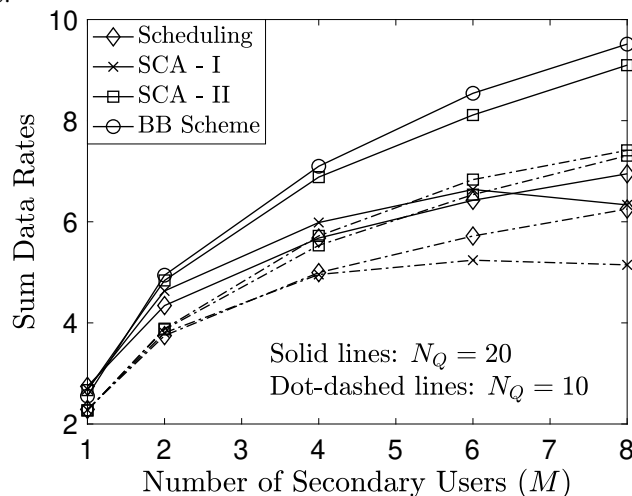


Fig. 5. Impact of the size of the beamsteering codebook on the performance of THz-NOMA transmission. $N = 10$, $K = 4$, $R_k = 1$ BPCU, and $r_S = 10$.

the performance of THz-NOMA is reduced when there are more primary users, which can be explained as follows. Unlike OFDMA subcarriers, the K spatial beams are not orthogonal bandwidth resources for the secondary users. In particular, these beams have been tailored to the primary users' channels in order to ensure that there is no inter-beam interference between the primary users. Because the secondary users' channels are different from the primary users', the secondary users still experience inter-beam interference. This inter-beam interference can cause two types of performance degradation. First, each secondary user can suffer more interference from the primary users, if K increases. Second, by increasing K , more secondary users are scheduled, which further increases interference in the network.

In Fig. 4, the impact of the number of antennas at the base station on the performance of THz-NOMA networks is studied. As can be seen from the figure, increasing the number of

antennas at the base station reduces the sum rate achieved by THz-NOMA transmission. This reduction is expected and can be explained in the following. Recall that the K beams, \mathbf{f}_k , are designed to match the primary users' channel vectors. By increasing N , both the users' channel vectors and the K spatial beams become more directional, which makes it more challenging for a secondary user to find a matching beam. This performance degradation can be mitigated if the beams are designed by taking both the primary and secondary users' channels into consideration.

Recall that in this paper, each analog beamforming vector, $\tilde{\mathbf{f}}_k$, is selected from a codebook with the limited size (N_Q). Fig. 5 is provided to show the impact of this important system parameter, N_Q , on the performance of THz-NOMA transmission. In particular, Fig. 5 shows that the performance gain of THz-NOMA is larger by using a smaller N_Q , which can be explained in the following. The value of N_Q decides the resolution of analog beamforming. For example, $N_Q \rightarrow \infty$ means the use of perfect analog beamforming and $\tilde{\mathbf{f}}_k$ will be perfectly matched to the channel vector of primary user \mathbf{U}_k^P . Therefore, using analog beamforming with finite resolution, i.e., N_Q is small, provides an opportunity that the analog beamforming vector, $\tilde{\mathbf{f}}_k$, might not be perfect for primary user \mathbf{U}_k^P but potentially ideal for some secondary users. This observation that analog beamforming with finite resolution is beneficial for the implementation of NOMA is also consistent to the findings previously reported in [25].

VI. CONCLUSIONS

This paper has considered the use of NOMA as an add-on in THz networks. In particular, it was assumed that there exists a legacy THz system, where spatial beams have been configured to serve legacy primary users. The aim of this paper was to investigate how these pre-configured spatial beams can be used to serve additional secondary users without degrading the performance of the legacy network. The considered beam and power allocation problem was first formulated as a mixed combinatorial non-convex optimization problem, and then solved by two methods, one based on the BB method and the other based on SCA. Both analytical and simulation results have been presented to demonstrate that these spatial beams can be used as a type of bandwidth resources to connect additional users and yield a significant throughput gain. Note that this gain is achieved without degrading the performance of the legacy system or acquiring additional spectrum. However, unlike conventional bandwidth resources, spatial beams are non-orthogonal resources, and hence it is important to study how to suppress inter-beam interference, which is an important direction for future research.

APPENDIX A
PROOF FOR LEMMA 1

The assumption that $\bar{R}_k^P \rightarrow 0$ means that $c_k \rightarrow -\infty$ and $b_{1k} \rightarrow -\infty$. Therefore, $P^{\max} = \min\{P^{\max}, -c_k, -b_{1k}\}$ and the constraints in (P2b), (P2c), and (P2d) are always satisfied. By using this assumption, problem P2 can be approximated at high SNR as follows:

$$\max_{\rho_k} \sum_{k=1}^K s_k \log \left(1 + \frac{h_k^S \rho_k^S}{\sum_{i=1, i \neq k}^K h_i^S s_i \rho_i^S + \rho^P \sum_{i=1, i \neq k}^K h_i^S} \right) \quad (\text{P20a})$$

$$\text{s.t.} \quad \sum_{k=1}^K s_k = 1, \quad s_k \in \{0, 1\}, \quad \sum_{k=1}^K \rho_k^S \leq P^{\max}. \quad (\text{P20b})$$

where ρ^P denotes the primary users' transmit power, and the notations, h_{1k}^S , ρ_{1k}^S , and s_{1k} , are simplified as h_k^S , ρ_k^S , and s_k , respectively.

For the considered special case, it is straightforward to show that the optimal solution of the greedy scheduling problem formulated in (P3) is simply given by $\rho_{k^*}^S = P^{\max}$, where $k^* = \arg \max\{h_k^S, 1 \leq k \leq 2\}$. Without loss of generality, assume that the secondary user's effective channel gains on the two beams are ordered as follows: $h_1^S > h_2^S$. Therefore, the key step to prove that problems P3 and P20 have the same optimal solution is to show that assuming that $\rho_1^S + \rho_2^S = \rho^S \leq \rho^{\max}$, the optimal solutions of ρ_1^S and ρ_2^S are ρ^S and 0, respectively. Once this step is established, it is straightforward to show that the optimal value of ρ^S is ρ^{\max} .

To simplify the proof, assume that $\alpha \rho^S$ is allocated to the first beam and the $(1 - \alpha) \rho^S$ is allocated to the second beam, $0 \leq \alpha \leq 1$, which means that the objective function is given by

$$f(\alpha) \triangleq \log \left(1 + \frac{h_1^S \alpha \rho^S}{h_2^S (1 - \alpha) \rho^S + h_2^S \rho^P} \right) + \log \left(1 + \frac{h_2^S (1 - \alpha) \rho^S}{h_1^S \alpha \rho^S + h_1^S \rho^P} \right), \quad (41)$$

where $0 \leq \alpha \leq 1$. The remainder of the proof is to show that the objective function in (41) is maximized by $\alpha = 1$, regardless of the choices of h_1^S , h_2^S , ρ^S and ρ^P .

Note that $f(\alpha)$ can be first rewritten as follows:

$$\begin{aligned} f(\alpha) &= \log \left(1 + \frac{h_1^S \alpha}{h_2^S (1 - \alpha) + \beta h_2^S} \right) + \log \left(1 + \frac{h_2^S (1 - \alpha)}{h_1^S \alpha + \beta h_1^S} \right) \\ &= \log \left(1 + x \frac{\alpha}{(1 + \beta - \alpha)} \right) + \log \left(1 + \frac{1}{x} \frac{(1 - \alpha)}{\alpha + \beta} \right), \end{aligned} \quad (42)$$

where $\beta = \frac{\rho^P}{\rho^S}$, $x = \frac{h_1^S}{h_2^S}$ and $x > 1$ because of the assumption that $h_1^S > h_2^S$.

It is important to point out that $f(\alpha)$ can be either convex or concave, depending on the choice of x and β , as shown in Fig. 6. However, $\alpha = 1$ always maximizes $f(\alpha)$, regardless the choices

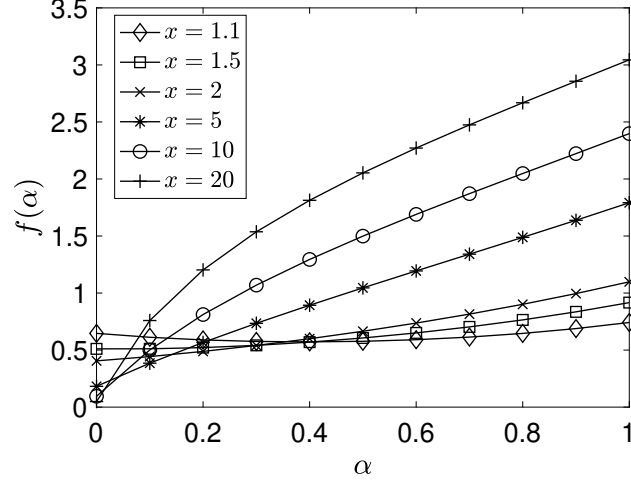


Fig. 6. Property of the function $f(\alpha)$ with $\beta = 1$.

of x and β . To show that $\alpha = 1$ maximizes $f(\alpha)$, the first order derivative of $f(\alpha)$ with respect to α is first obtained as follows:

$$\begin{aligned} \frac{df(\alpha)}{d\alpha} &= \frac{x-1}{1+\beta+(x-1)\alpha} - \frac{1}{\alpha-(1+\beta)} + \frac{x-1}{(x-1)\alpha+x\beta+1} - \frac{1}{\alpha+\beta} \\ &= \frac{-(1+\beta)\left((x+1)\alpha^2 + \left(x\left(\beta + \frac{x\beta+1}{x-1}\right) + \frac{1+\beta}{x-1} - (1+\beta)\right)\alpha + \frac{x^2\beta^2+x\beta}{x-1} - \frac{(1+\beta)^2}{x-1}\right)}{\left(\alpha + \frac{x\beta+1}{x-1}\right)(\alpha+\beta)\left(\alpha + \frac{1+\beta}{x-1}\right)(\alpha-(1+\beta))(x-1)}. \end{aligned} \quad (43)$$

Note that the quadratic function in the numerator of (43) is concave since $x > 1$, $\alpha \leq 1$ and $\beta \geq 0$.

As shown in Fig. 6, $f(\alpha)$ can be convex or concave, which makes the proof challenging. Interestingly, $\frac{df(\alpha)}{d\alpha}$ is always positive at $\alpha = 1$ for any choices of x and β , as shown in the following:

$$\begin{aligned} \left. \frac{df(\alpha)}{d\alpha} \right|_{\alpha=1} &= \frac{(x+1) + x\left(\beta + \frac{x\beta+1}{x-1}\right) + \frac{1+\beta}{x-1} - (1+\beta) + \frac{x^2\beta^2+x\beta}{x-1} - \frac{(1+\beta)^2}{x-1}}{\left(1 + \frac{x\beta+1}{x-1}\right)\left(1 + \frac{1+\beta}{x-1}\right)\beta(x-1)} \\ &= \frac{(1+\beta)^2x^2 - \beta x - \beta^2}{\left(1 + \frac{x\beta+1}{x-1}\right)\left(1 + \frac{1+\beta}{x-1}\right)\beta(x-1)^2}. \end{aligned} \quad (44)$$

In order to show $\left. \frac{df(\alpha)}{d\alpha} \right|_{\alpha=1} > 0$, it is sufficient to show that the numerator of (44), defined as $f_\beta(x) = (1+\beta)^2x^2 - \beta x - \beta^2$, is positive for $x > 1$ and $\beta \geq 0$. Note that the two roots of $f_\beta(x) = 0$ are $\frac{\beta \pm \sqrt{\beta^2 + 4\beta^2(1+\beta)^2}}{2(1+\beta)^2}$. Therefore, the proof can be complete by showing that the positive root $\frac{\beta + \sqrt{\beta^2 + 4\beta^2(1+\beta)^2}}{2(1+\beta)^2} < 1$, which can be established due to the equivalence between the following two inequalities:

$$\sqrt{\beta^2 + 4\beta^2(1+\beta)^2} < 2(1+\beta)^2 - \beta \iff \beta < (1+2\beta). \quad (45)$$

The fact that $\left. \frac{df(\alpha)}{d\alpha} \right|_{\alpha=1} > 0$ is important because it shows that $f(\alpha)$ is an increasing function at $\alpha = 1$. Denote the two roots of the quadratic function in the numerator of (43) by r_1 and r_2 . Without loss of generality, assume $r_1 \leq r_2$. The use of $\left. \frac{df(\alpha)}{d\alpha} \right|_{\alpha=1} > 0$ leads to the conclusion that $r_1 \leq 1 \leq r_2$. Another important fact is that $f(0) < f(1)$, since scheduling the weak user results in a smaller data rate compared to the case with the strong user scheduled. By using the two facts, the proof for $\max f(\alpha) = f(1)$ can be established as follows:

1) If $r_1 < 0$: $\left. \frac{df(\alpha)}{d\alpha} \right|_{\alpha=1}$ is positive for $0 \leq \alpha \leq 1$, which means that $f(\alpha)$ is an increasing function for $0 \leq \alpha \leq 1$. In Fig. 6, the curves with $x = 5$, $x = 10$, $x = 20$, $x = 1.5$ and $x = 2$ belong to this case. Therefore, $\alpha = 1$ can maximize the objective function $f(\alpha)$.

2) If $r_1 \geq 0$: $\left. \frac{df(\alpha)}{d\alpha} \right|_{\alpha=1}$ is first non-positive for $0 \leq \alpha < r_1$, and then becomes positive for $r_1 \leq \alpha \leq 1$, which means that $f(\alpha)$ is a non-increasing function for $0 \leq \alpha < r_1$, and then becomes an increasing function $r_1 \leq \alpha \leq 1$. In Fig. 6, the curve with $x = 1.1$ belongs to this case. Furthermore, by using the fact that $f(0) < f(1)$, $\alpha = 1$ can still maximize the objective function $f(\alpha)$ in this case.

In summary, $\alpha = 1$ can always maximize the objective function, and hence the proof for the lemma is complete.

APPENDIX B

PROOF FOR LEMMA 2

Recall that the aim of the tightening procedure is to find the maximal value of $\tilde{x}_{\max,p}$, which can be achieved by solving the following optimization problem:

$$\max_{\mathbf{y}} \frac{\mathbf{c}_p^T \mathbf{y}}{\mathbf{d}_p^T \mathbf{R} \mathbf{y} + t_{\mathbf{S}_{p1}} \mathbf{S}_{p2}} \quad (\text{P21a})$$

$$\text{s.t. } \mathbf{A}_p \mathbf{y} \leq \mathbf{b}_p \quad (\text{P21b})$$

$$\mathbf{A}_p^E \mathbf{y} = \mathbf{b}_p^E. \quad (\text{P21c})$$

It is important to point out that the optimization variable vector, \mathbf{y} , contains $|\mathcal{S}|$ elements, whereas (P21c) contains $(|\mathcal{S}| - 1)$ equality constraints. This important observation can be used to reduce the number of optimization variables from $|\mathcal{S}|$ to one only, as shown in the following.

Without loss of generality, the constraint in (P21c) can be expressed as follows:

$$\mathbf{a}_{p,p}^E y_p + \tilde{\mathbf{A}}_p^E \tilde{\mathbf{y}}_p = \mathbf{b}_p^E. \quad (46)$$

Assuming that $\tilde{\mathbf{A}}_p^E$ is an invertible matrix, the $(|\mathcal{S}| - 1)$ optimization variables in $\tilde{\mathbf{y}}_p$ can be expressed as the following functions of y_p :

$$\tilde{\mathbf{y}}_p = (\tilde{\mathbf{A}}_p^E)^{-1}(\mathbf{b}_p^E - \mathbf{a}_{p,p}^E y_p). \quad (47)$$

By using the fact that \mathbf{c}_p is a $|\mathcal{S}| \times 1$ all-zero vector except its p -th element, denoted by $c_{p,p}$, problem P21 can be recast as the following optimization problem:

$$\max_{\mathbf{y}} \frac{c_{p,p} y_p}{\bar{d}_p y_p + \tilde{\mathbf{d}} \tilde{\mathbf{y}}_p + t_{\mathbf{s}_{p1}} \mathbf{s}_{p2}} \quad (P22a)$$

$$\text{s.t.} \quad \mathbf{a}_{p,p} y_p + \tilde{\mathbf{A}}_p \tilde{\mathbf{y}}_p \leq \mathbf{b}_p, \quad (P22b)$$

where $\bar{\mathbf{d}} = \mathbf{d}_p^T \mathbf{R}$, \bar{d}_p denotes the p -th element of $\bar{\mathbf{d}}$, and $\tilde{\mathbf{d}}$ is obtained from $\bar{\mathbf{d}}$ by removing \bar{d}_p .

Furthermore, by using the fact that $(|\mathcal{S}| - 1)$ optimization variables in $\tilde{\mathbf{y}}_p$ can be expressed as a function of y_p , constraint (P22b) can be rewritten as follows:

$$\mathbf{a}_{p,p} y_p + \tilde{\mathbf{A}}_p (\tilde{\mathbf{A}}_p^E)^{-1} (\mathbf{b}_p^E - \mathbf{a}_{p,p}^E y_p) \leq \mathbf{b}_p, \quad (48)$$

which can be further re-written as follows:

$$(\mathbf{a}_{p,p} - \tilde{\mathbf{A}}_p (\tilde{\mathbf{A}}_p^E)^{-1} \mathbf{a}_{p,p}^E) y_p \leq (\mathbf{b}_p - \tilde{\mathbf{A}}_p (\tilde{\mathbf{A}}_p^E)^{-1} \mathbf{b}_p^E). \quad (49)$$

By using (49), problem P17 can be expressed as the following simple optimization problem:

$$\max_{\mathbf{y}} \frac{c_{p,p} y_p}{\bar{d}_p y_p + \tilde{\mathbf{d}} \tilde{\mathbf{y}}_p + t_{jk}} \quad (P23a)$$

$$\text{s.t.} \quad y_p \leq \mathbf{a}_{\text{sign}} \odot (\mathbf{b}_p - \tilde{\mathbf{A}}_p (\tilde{\mathbf{A}}_p^E)^{-1} \mathbf{b}_p^E) ./ (\mathbf{a}_{p,p} - \tilde{\mathbf{A}}_p (\tilde{\mathbf{A}}_p^E)^{-1} \mathbf{a}_{p,p}^E). \quad (P23b)$$

Note that the following function, $\frac{ax}{bx+1}$, is a monotonically increasing function of x for $x \geq 0$, $a \geq 0$ and $b \geq 0$. Therefore, it is straightforward to show that the optimal solution of problem P23 can be obtained as shown in the lemma.

REFERENCES

- [1] X. You, C. Wang, J. Huang *et al.*, "Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts," *Sci. China Inf. Sci.*, vol. 64, no. 110301, pp. 1–74, Feb. 2021.
- [2] Z. Zhang, Y. Xiao, Z. Ma, M. Xiao, Z. Ding, X. Lei, G. K. Karagiannidis, and P. Fan, "6G wireless networks: Vision, requirements, architecture, and key technologies," *IEEE Veh. Tech. Mag.*, vol. 14, no. 3, pp. 28–41, Jul. 2019.
- [3] H.-J. Song and T. Nagatsuma, "Present and future of terahertz communications," *IEEE Trans. THz Sci. Technol.*, vol. 1, no. 1, pp. 256–263, Sept. 2011.
- [4] X. Tong, B. Chang, Z. Meng, G. Zhao, and Z. Chen, "Calculating terahertz channel capacity under beam misalignment and user mobility," *IEEE Wireless Commun. Lett.*, vol. 11, no. 2, pp. 348–351, Feb. 2022.

- [5] A.-A. A. Boulogeorgos, E. N. Papatirou, and A. Alexiou, "Analytical performance assessment of THz wireless systems," *IEEE Access*, vol. 7, pp. 11 436–11 453, 2019.
- [6] N. Olson, J. G. Andrews, and R. W. Heath, "Coverage in Terahertz cellular networks with imperfect beam alignment," in *Proc. IEEE Globecom*, Madrid, Spain, Dec. 2021.
- [7] M. Vaezi, Z. Ding, and H. V. Poor, *Multiple Access Techniques for 5G Wireless Networks and Beyond*. Springer International Publishing, 2019.
- [8] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. 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.
- [9] H. Sameddeen, A. Abdallah, M. M. Mansour, M.-S. Alouini, and T. Y. Al-Naffouri, "Terahertz-band MIMO-NOMA: Adaptive superposition coding and subspace detection," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 2628–2644, 2021.
- [10] Z. Ding and H. V. Poor, "Design of THz-NOMA in the presence of beam misalignment," *IEEE Commun. Lett.*, to appear in 2022.
- [11] O. Maraqa, A. S. Rajasekaran, H. U. Sokun, S. Al-Ahmadi, H. Yanikomeroglu, and S. M. Sait, "Energy-efficient coverage enhancement of indoor THz-MISO systems: An FD-NOMA approach," in *Proc. IEEE Int. Symposium on Personal, Indoor and Mobile Radio Commun.*, Helsinki, Finland, Sept. 2021.
- [12] H. Zhang, H. Zhang, W. Liu, K. Long, J. Dong, and V. C. M. Leung, "Energy efficient user clustering, hybrid precoding and power optimization in terahertz MIMO-NOMA systems," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 9, pp. 2074–2085, Sept. 2020.
- [13] H. Zhang, Y. Duan, K. Long, and V. C. M. Leung, "Energy efficient resource allocation in terahertz downlink NOMA systems," *IEEE Trans. Commun.*, vol. 69, no. 2, pp. 1375–1384, Feb. 2021.
- [14] S. Bani Melhem and H. Tabassum, "User pairing and outage analysis in multi-carrier NOMA-THz networks," *IEEE Trans. Veh. Tech.*, to appear in 2022.
- [15] X. Xu, Q. Chen, X. Mu, Y. Liu, and H. Jiang, "Graph-embedded multi-agent learning for smart reconfigurable THz MIMO-NOMA networks," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, pp. 259–275, Jan. 2022.
- [16] Z. Ding, "NOMA beamforming in SDMA networks: Riding on existing beams or forming new ones?" *IEEE Commun. Lett.*, vol. 26, no. 4, pp. 868–871, Apr. 2022.
- [17] C. Y. Wong, R. Cheng, K. Lataief, and R. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.
- [18] J. Cui, Y. Liu, Z. Ding, P. Fan, and A. Nallanathan, "Optimal user scheduling and power allocation for millimeter wave NOMA systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1502–1517, 2018.
- [19] P. C. Weeraddana, M. Codreanu, M. Latva-aho, and A. Ephremides, "Weighted sum-rate maximization for a set of interfering links via branch and bound," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3977–3996, Aug. 2011.
- [20] W.-C. Li, T.-H. Chang, C. Lin, and C.-Y. Chi, "Coordinated beamforming for multiuser miso interference channel under rate outage constraints," *IEEE Transactions on Signal Processing*, vol. 61, no. 5, pp. 1087–1103, Mar. 2013.
- [21] Y. Xu, C. Shen, Z. Ding, X. Sun, S. Yan, G. Zhu, and Z. Zhong, "Joint beamforming and power-splitting control in downlink cooperative SWIPT NOMA systems," *IEEE Trans. Signal Process.*, vol. 65, no. 18, pp. 4874–4886, Sept. 2017.
- [22] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.
- [23] Y. Zou, W. Rave, and G. Fettweis, "Analog beamsteering for flexible hybrid beamforming design in mmWave communications," in *Proc. European Conference on Networks and Communications (EuCNC)*, Athens, Greece, Jun. 2016.
- [24] Y. Sun, D. Xu, D. W. K. Ng, L. Dai, and R. Schober, "Optimal 3D-trajectory design and resource allocation for solar-powered UAV communication systems," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4281–4298, Jun. 2019.
- [25] Z. Ding, L. Dai, R. Schober, and H. V. Poor, "NOMA meets finite resolution analog beamforming in massive MIMO and millimeter-wave networks," *IEEE Commun. Lett.*, vol. 21, no. 8, pp. 1879–1882, Aug. 2017.