User Fairness Non-Orthogonal Multiple Access (NOMA) for Millimeter-Wave Communications With Analog Beamforming

Zhenyu Xiao, *Senior Member, IEEE*, Lipeng Zhu, Zhen Gao, *Member, IEEE*, Dapeng Oliver Wu, *Fellow, IEEE*, and Xiang-Gen Xia, *Fellow, IEEE*

***Abstract*— The integration of non-orthogonal multiple access in millimeter-Wave communications (mm Wave-NOMA) can significantly improve the spectrum efficiency and increase the number of users in the fifth-generation (5G) mobile communication and beyond. In this paper, we consider a downlink mm Wave-NOMA cellular system, where the base station is mounted with an analog beamforming phased array, and multiple users are served in the same time-frequency resource block. To guarantee user fairness, we formulate joint beamforming and power allocation problem to maximize the minimal achievable rate among the users, i.e., we adopt the max–min fairness. As the problem is difficult to solve due to the non-convex formulation and high dimension of the optimization variables, we propose a sub-optimal solution, which makes use of the spatial sparsity in the angle domain of the mm Wave channel. In the solution, the closed-form optimal power allocation is obtained first, which reduces the joint optimization problem into an equivalent beamforming problem. Then, an appropriate beamforming vector is designed. The simulation results show that the proposed solution can achieve a near-upper-bound performance in terms of achievable rate, which is significantly better than that of the conventional mm Wave orthogonal multiple access (mm Wave-OMA) system.**

***Index Terms*— Millimeter-wave communications, non-orthogonal multiple access, mm Wave-NOMA, user fairness, analog beamforming, power allocation.**

I. Introduction

With the rapid growth of mobile data traffic, higher data rate is an insistent requirement in the fifth generation (5G) of mobile communication and beyond [1], [2]. Millimeter-Wave (mm Wave) communications, with frequency ranging from 30-300 GHz, provides abundant spectrum resources and is perceived as a candidate key technology for 5G and beyond [1], [3], [4]. In addition to the large amount of bandwidth, the mm Wave-band signal has a shorter wavelength compared with the traditional microwave-band signal, which makes it possible to equip a large antenna array in a small area. Considerable beam gain can be obtained to overcome the high propagation loss in the mm Wave-band [4].

Although more spectrum resources are available in the mm Wave band, multiple access is still an important issue to increase the spectrum efficiency and the number of users/devices to support 5G Internet of Things (IoT). Non-orthogonal multiple access (NOMA), considered as another candidate technology for 5G, has drawn widespread attention in both academia and industry [5]–[12]. Different from the conventional orthogonal multiple access (OMA) schemes, NOMA serves multiple users in a single resource block (time/frenquency) and distinguishes them in power domain or code domain, which can be mainly classified into power-domain NOMA and code-domain NOMA, respectively. For power-domain NOMA, the signals are superimposed with different transmission powers at the transmitter, and successive interference cancellation (SIC) is required at the receivers to remove the multi-user interference [6]–[11]. For code-domain NOMA, the users are allocated with sparse sequences or non-orthogonal low cross-correlation sequences, e.g., low-density spreading (LDS), sparse code multiple access (SCMA), multi-user shared access (MUSA), pattern division multiple access (PDMA), and so on [9], [11], [13].

To make full use of the spectrum resource, we investigate power-domain NOMA in mm Wave communications (mm Wave-NOMA) in this paper, which has the unique advantage. Conventionally, the number of data streams is no larger than that of radio frequency (RF) chains. Subject to hardware cost and power consumption, the number of RF chains is usually limited in mm Wave communications, which means that the number of data streams is limited. Thus, NOMA can be introduced to support more data streams or more users in mm Wave communications [14]. On the other hand, flexible mm Wave beamforming provides an additional degree of freedom for NOMA user grouping and performance improvement [5], [12], [14].

The combination of the two candidate technologies for 5G has been preliminarily explored in several literatures. In [15], the coexistence of NOMA and mm Wave communications was

Manuscript received August 3, 2018; revised January 15, 2019; accepted April 19, 2019. Date of publication May 6, 2019; date of current version July 10, 2019. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61571025, Grant 91538204, Grant 61827901, Grant 91738301, and Grant 61701027, and in part by the Beijing Natural Science Foundation under Grant 4182055. The associate editor coordinating the review of this paper and approving it for publication was M. Payaró. *(Corresponding author: Zhenyu Xiao.)*

Z. Xiao and L. Zhu are with the School of Electronic and Information Engineering, Beihang University, Beijing 100191, China (e-mail: xiaozy@buaa.edu.cn).

Z. Gao is with the Advanced Research Institute of Multidisciplinary Science, Beijing Institute of Technology, Beijing 100081, China.

D. O. Wu is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA.

X.-G. Xia is with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2019.2913844

considered, where random beamforming was used in order to reduce the system overhead. The results demonstrated that the combination of NOMA and mm Wave communications yields significant gains in terms of sum rates and outage probabilities, compared with the conventional mm Wave-OMA systems. In [16], the new concept of beamspace multiple-input multiple-output NOMA (MIMO-NOMA) with a lens-array hybrid beamforming structure was proposed to use multi-beam to serve multiple NOMA users with arbitrary locations. With this method, the number of supported users can be larger than the number of RF chains in the same time-frequency resource block. Beamforming, user selection and power allocation were considered for mm Wave-NOMA networks in [17], where random beamforming was adopted first. Then a power allocation algorithm that leverages the branch and bound (BB) technique and a low complexity user selection algorithm based on matching theory were proposed. A NOMA based hybrid beamforming design was proposed in [18], where a user pairing algorithm was proposed first and then the hybrid beamforming and power allocation algorithm was proposed to maximize the sum achievable rate. In [19], the NOMA-mm Wave-massive-MIMO system model and a simplified mm Wave channel model were proposed. Whereafter, theoretical analysis on the achievable rate was considered in both the noise-dominated low-SNR regime and the interference-dominated high-SNR regime. To further improve the data rate, power allocation and beamforming were jointly explored in [20] and [21] for a 2-user downlink and uplink mm Wave-NOMA scenario, respectively, where the key technique is the multi-directional beamforming design with a constant-modulus (CM) phased array. In [22], a multi-beam NOMA scheme with subarray techniques for hybrid mm Wave systems was proposed, and the user grouping, antenna and power allocation were designed to maximize the sum rate.[1]

Different from these works [15]–[21], we consider user fairness for downlink mm Wave-NOMA networks in this paper. To improve the overall data rate, we maximize the minimal achievable rate among multiple users, i.e., we adopt the max-min fairness.[2] Due to the requirement of low hardware cost and power consumption, an analog beamforming structure with a single RF chain is utilized, where both implementations of single phase shifter (SPS) and double phase shifter (DPS) are considered [26], [27]. It is worthwhile pointing out that although there are several literatures exploring the user fairness for NOMA and MIMO-NOMA, the key feature of mm Wave communications, i.e., analog beamforming, was not considered [25], [28]–[31]. Thus, the proposed approaches in [25], [28]–[31] cannot be directly used to solve the problem for user fairness mm Wave-NOMA considered in this paper. In our formulated problem, power allocation and beamforming are jointly optimized. As the problem is non-convex and the dimension of the optimization variables is large, it is difficult to solve this problem with the existing optimization tools. To this end, we solve this problem with two stages and obtain a sub-optimal solution. In the first stage, we obtain closed-form optimal power allocation with an arbitrary fixed beamforming vector, which reduces the joint optimization problem into an equivalent beamforming problem. Then, in the second stage, we propose an appropriate beamforming algorithm utilizing the spatial sparsity in the angle domain of the mm Wave channel. Finally, we verify the performance of the proposed joint beamforming and power allocation method for user fairness mm Wave-NOMA by simulations. The results show that the proposed solution can achieve a near-upper-bound performance in terms of achievable rate, which is significantly better than that of the conventional mm Wave-OMA system.

The rest of the paper is organized as follows. In Section II, we present the system model and formulate the problem. In Section III, we propose the solution. In Section IV, simulation results are given to demonstrate the performance of the proposed solution, and the paper is concluded finally in Section V.

Symbol Notation: a and $\mathbf{a}$ denote a scalar variable and a vector, respectively. $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ denote transpose and conjugate transpose, respectively. $|\cdot|$ and $\|\cdot\|$ denote the absolute value and Euclidean norm, respectively. $\mathbb{E}(\cdot)$ denotes the expectation operation. $[\mathbf{a}]_i$ denotes the i-th entry of $\mathbf{a}$. $\mathbb{C}^N$ denotes an N-dimension linear space in complex domain.

II. System Model and Problem Formulation

A. System Model

In this paper, we consider a downlink mm Wave communications system. As shown in Fig. 1, the base station (BS) is equipped with a single RF chain and an N-antenna phased array.[3] K users with a single antenna are served in the same resource block. Each antenna is driven by the power amplifier (PA) and phase shifter (PS). With only one RF chain at the BS, we only need to consider the analog beamforming, which has a low hardware complexity and it can also be used in the hybrid beamforming structure.

The BS transmits a signal s_k to user k $(k = 1, 2, \cdots, K)$ with transmission power p_k, where $\mathbb{E}(|s_k|^2) = 1$. The total transmission power of the BS is P. The received signal for user k is

$$y_k = \mathbf{h}_k^{\mathrm{H}} \mathbf{w} \sum_{k=1}^{K} \sqrt{p_k} s_k + n_k, \tag{1}$$

where $\mathbf{h}_k$ is the channel response vector between the BS and user k. $\mathbf{w}$ is the antenna weight vector (AWV), i.e., analog

[1] In [22], multiple beams in the analog domain are obtained by using different subarrays to steer to different users. As the subarray technique results in a wider beamwidth in the angle domain [23], [24], this method may not fully explore the capability of the beamforming. In particular, when the number of users increases, the beam gains may decrease rapidly. In this paper, we use the optimization approach to improve the multi-beam gains and maximize the minimal rate by jointly designing the beamforming and power allocation.

[2] We adopt the max-min fairness because it is a typical and extensively used fairness rule in NOMA [11], [25]. Besides the max-min fairness, there are also other fairness rules in NOMA, like proportional fairness, etc. [11].

[3] Note that in hybrid beamforming, the design of an analog precoder has also been widely studied [16]–[18], [22], where the precoder is usually a matrix with constant modulus. Each column of the precoding matrix is in fact an analog beamforming vector. Hence, for the same optimization object, the proposed approach in this paper can also be used when designing an analog precoder of hybrid beamforming. However, due to the existence of digital precoding, hybrid beamforming is usually much more complicated than analog beamforming. It would be interesting to study hybrid beamforming for the problem in this paper, which will be scheduled in our next work.

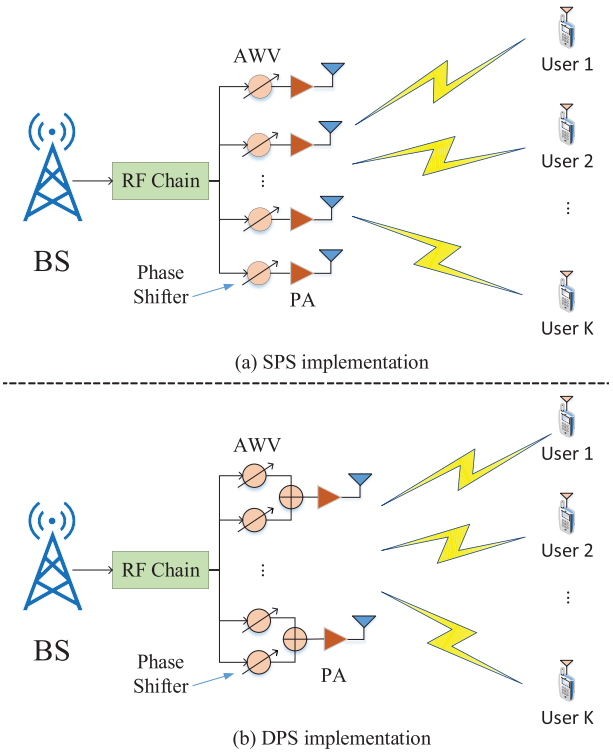

Fig. 1. Illustration of a mm Wave mobile cell, where one BS with N antennas serves multiple users with one single antenna.

beamforming vector, and n_k denotes the Gaussian white noise at user k with power σ^2.

The conventional SPS implementation at the BS is shown in Fig. 1(a), where each antenna branch has a single PS. The elements of the AWV are complex numbers, whose modulus and phase are controlled by the PA and PS, respectively. In general, all the PAs have the same scaling factor to reduce the hardware complexity. Thus, the AWV of the SPS implementation has CM elements, which is denoted by

$$|[\mathbf{w}]_i| = \frac{1}{\sqrt{N}},\ i = 1, 2, ..., N. \tag{2}$$

Since the above constraint for the SPS implementation is non-convex, it results in high computational complexity of designing the AWV if the design space has a large dimension. Although some codebooks have been designed for the SPS implementation, there is still a tradeoff between the beamforming performance and the computational complexity [23], [24]. To address this problem, a new implementation named DPS was proposed in [26], [27], which is shown in Fig. 1(b). For the DPS implementation, each antenna is driven by the summation of the two independent PSs. Although the modulus of each PS is constant, the phases of two PSs can be adjusted to achieve different modulus in each antenna branch. Thus, the modulus constraint is relaxed to

$$|[\mathbf{w}]_i| \leq \frac{2}{\sqrt{N}},\ i = 1, 2, ..., N. \tag{3}$$

By doubling the number of PSs, the new constraint becomes convex and therefore makes it more tractable to develop low-complexity design approaches. With this implementation, it is possible to achieve a better beamforming performance with lower computational complexity.

The channel between the BS and user k is a mm Wave channel.[4] Subject to the limited scattering in mm Wave-band, multipath is mainly caused by reflection. As the number of the multipath components (MPCs) is small in general, the mm Wave channel has directionality and appears spatial sparsity in the angle domain [23], [32]–[36]. Different MPCs have different angles of departure (AoDs). Without loss of generality, we adopt the directional mm Wave channel model assuming a uniform linear array (ULA) with a half-wavelength antenna space. Then a mm Wave channel can be expressed as [23], [32]–[36]

$$\mathbf{h}_k = \sum_{\ell=1}^{L_k} \lambda_{k,\ell} \mathbf{a}(N, \Omega_{k,\ell}). \tag{4}$$

where $\lambda_{k,\ell}$, $\Omega_{k,\ell}$ are the complex coefficient and cos(AoD) of the ℓ-th MPC of the channel vector for user k, respectively. We have $\sum_{l=1}^{L_k} \mathbb{E}(|\lambda_{k,\ell}|^2) \propto \frac{1}{d_k^2}$, where d_k is the distance between the BS and user k [15], [17]. L_k is the total number of MPCs for user k, $\mathbf{a}(\cdot)$ is a steering vector function defined as

$$\mathbf{a}(N, \Omega) = [e^{j\pi 0\Omega}, e^{j\pi 1\Omega}, e^{j\pi 2\Omega}, \cdots, e^{j\pi(N-1)\Omega}]^{\mathrm{T}}, \tag{5}$$

which depends on the array geometry. Let $\theta_{k,\ell}$ denote the real AoD of the ℓ-th MPC for user k, then we have $\Omega_{k,\ell} = \cos(\theta_{k,\ell})$. Therefore, $\Omega_{k,\ell}$ is within the range $[-1, 1]$.

In general, the optimal decoding order of NOMA is the increasing order of the effective channel gains [7], [13], i.e., $\left|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}\right|^2$. However, we cannot determine the order of the effective channel gains before beamforming. For simplicity, we utilize the increasing order of uses' channel gains as the decoding order. We will illustrate the rational of selecting the increasing-channel-gain decoding order in Section III-C, and verify that it can achieve near optimal performance by simulations. Without loss of generality, we assume $\|\mathbf{h}_1\| \geq \|\mathbf{h}_2\| \geq \cdots \geq \|\mathbf{h}_K\|$. Therefore, user k can decode s_n $(k{+}1 \leq n \leq K)$ and then remove them from the received signal in a successive manner. The signals for user m $(1 \leq m \leq k-1)$ are treated as noise. Thus, the achievable rate of user k is given by

$$R_k = \log_2\left(1 + \frac{\left|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}\right|^2 p_k}{\left|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}\right|^2 \sum_{m=1}^{k-1} p_m + \sigma^2}\right). \tag{6}$$

B. Problem Formulation

As aforementioned, both beamforming and power allocation have an important effect on the performance of the mm Wave-NOMA system. To improve the overall data rate and guarantee user fairness, we formulate a problem to maximize the minimal achievable rate (the max-min fairness) among the K users

[4] As we focus on the resource allocation for mm Wave-NOMA, the channel estimation is beyond the scope of this paper. Thus, we assume the channel state information (CSI) is known by the BS. In practice, many efficient methods with low complexity [32], [33] and [24] can be referred to estimating the mm Wave channel.

in this paper, where beamforming and power allocation are jointly optimized. The problem is formulated as

$$\begin{aligned}&\underset{\{p_k\},\mathbf{w}}{\text{Max}} \ \min_k\{R_k\}\\ &\text{s.t. } C_1:\ p_k \geq 0, \quad k=1,2,\cdots,K\\ &\qquad C_2:\ \sum_{k=1}^{K} p_k \leq P,\\ &\qquad C_3:\ \|\mathbf{w}\| \leq \mathbf{1},\\ &\qquad C_4:\ |[\mathbf{w}]_i| = \frac{1}{\sqrt{N}} \text{ or } |[\mathbf{w}]_i| \leq \frac{2}{\sqrt{N}},\ i=1,2,...,N\end{aligned} \tag{7}$$

where R_k denotes the achievable rate of user k as defined in (6) and $\min_k\{R_k\}$ is the minimal achievable rate among the K served users. The constraint C_1 indicates that the power allocation to each user should be nonnegative. C_2 is the transmission power constraint, where P is the total transmission power. C_3 is the unit norm constraint on the AWV. As the modulus constraints on the AWV for different PS implementations are different, we distinguish them in the constraint C_4. It will be shown later, with the same computational complexity, the DPS implementation can achieve a better performance compared with the SPS implementation.

The above problem is challenging, not only due to the non-convex formulation, but also due to that the variables to be optimized are entangled with each other. It is computationally prohibitive to directly search the optimal solution, because the dimension of the optimization variables is $N+K$, which is large in general. Next, we will propose a sub-optimal solution with promising performance but low computational complexity.

III. Solution of the Problem

As the modulus constraints for SPS and DPS implementations are different, we first solve the problem without considering the constraint C_4. As thus, Problem (7) is simplified as

$$\begin{aligned}&\underset{\{p_k\},\mathbf{w}}{\text{Max}} \ \min_k\{R_k\}\\ &\text{s.t. } C_1:\ p_k \geq 0, \quad k=1,2,\cdots,K\\ &\qquad C_2:\ \sum_{k=1}^{K} p_k \leq P,\\ &\qquad C_3:\ \|\mathbf{w}\| \leq \mathbf{1}.\end{aligned} \tag{8}$$

We will solve Problem (8) first, and then particularly consider the modulus constraints in Section III-D.

Problem (8) is still difficult due to the non-convex formulation, so we propose a sub-optimal solution with two stages. In the first stage, we obtain the closed-form optimal power allocation with an arbitrary fixed AWV. Then, in the second stage, we propose an appropriate beamforming algorithm utilizing the angle-domain spatial sparsity of the mm Wave channel. It is worth pointing out that a closed-form solution of the optimal power allocation obtained in the first stage is a function of the AWV. Thus, we can substitute it to Problem (8) and solve the beamforming problem in the second stage without loss of optimality.

A. Optimal Power Allocation With an Arbitrary Fixed AWV

First, we introduce a variable to simplify Problem (8). Denote the minimal achievable rate among the K users as r. Then Problem (8) can be re-written as

$$\begin{aligned}&\underset{\{p_k\},\mathbf{w},r}{\text{Max}} \ r\\ &\text{s.t. } C_0:\ R_k \geq r, \quad k=1,2,\cdots,K\\ &\qquad C_1:\ p_k \geq 0, \quad k=1,2,\cdots,K\\ &\qquad C_2:\ \sum_{k=1}^{K} p_k \leq P,\\ &\qquad C_3:\ \|\mathbf{w}\| \leq \mathbf{1},\end{aligned} \tag{9}$$

where the constraints $C_0 : R_k \geq r,\ (k=1,2,\cdots,K)$ are necessary and sufficient conditions of the fact that r is the minimal achievable rate among the served users. On one hand, as r is the minimal rate, the achievable rate of each user should be no less than r. On the other hand, there is at least one user, whose achievable rate R_{k_m} is equal to r; otherwise we can always improve r to minish the gap between R_{k_m} and r.

We give the following Theorem to obtain the optimal solution of power allocation of Problem (9) with an arbitrary fixed AWV.

Theorem 1: Given an arbitrary fixed $\mathbf{w_0}$, the optimal power allocation of Problem (9) is

$$\begin{cases} p_1 = \eta\dfrac{\sigma^2}{\left|\mathbf{h}_1^{\mathrm{H}}\mathbf{w}_0\right|^2},\\ p_2 = \eta(p_1 + \dfrac{\sigma^2}{\left|\mathbf{h}_2^{\mathrm{H}}\mathbf{w}_0\right|^2}),\\ \quad\vdots\\ p_K = \eta(\sum\limits_{m=1}^{K-1} p_m + \dfrac{\sigma^2}{\left|\mathbf{h}_K^{\mathrm{H}}\mathbf{w}_0\right|^2}),\end{cases} \tag{10}$$

where $\eta = 2^r - 1$, and with the optimal power allocation, $R_k = r\ (k=1,2,\cdots,K)$.

Before proving Theorem 1, we give Lemma 1 for the summation of the optimal power allocation in (10), which is a function of η.

Lemma 1: The summation of power allocation in (10) is

$$g(\eta) \triangleq \sum_{k=1}^{K} p_k = \sum_{k=1}^{K} \frac{\eta(1+\eta)^{K-k}\sigma^2}{\left|\mathbf{h}_K^{\mathrm{H}}\mathbf{w}_0\right|^2}. \tag{11}$$

Proof: We prove Lemma 1 with mathematical induction. When $m=1$, (11) is easy to verify

$$p_1 = \eta\frac{\sigma^2}{\left|\mathbf{h}_1^{\mathrm{H}}\mathbf{w}_0\right|^2}. \tag{12}$$

When $m=n\ (n\geq 1)$, assume that

$$\sum_{k=1}^{n} p_k = \sum_{k=1}^{n} \frac{\eta(1+\eta)^{n-k}\sigma^2}{\left|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_0\right|^2}. \tag{13}$$

When $m = n+1$, based on (13), we have

$$\begin{aligned}\sum_{k=1}^{n+1} p_k &= \sum_{k=1}^{n} p_k + \eta\Big(\sum_{k=1}^{n} p_k + \frac{\sigma^2}{\left|\mathbf{h}_n^{\mathrm{H}}\mathbf{w}_0\right|^2}\Big) \\ &= (1+\eta)\sum_{k=1}^{n} p_k + \eta\frac{\sigma^2}{\left|\mathbf{h}_n^{\mathrm{H}}\mathbf{w}_0\right|^2} \\ &= (1+\eta)\sum_{k=1}^{n}\frac{\eta(1+\eta)^{n-k}\sigma^2}{\left|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_0\right|^2} + \eta\frac{\sigma^2}{\left|\mathbf{h}_n^{\mathrm{H}}\mathbf{w}_0\right|^2} \\ &= \sum_{k=1}^{n+1}\frac{\eta(1+\eta)^{n+1-k}\sigma^2}{\left|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_0\right|^2}. \end{aligned} \tag{14}$$

Finally, we can conclude that (11) is true. ■

Based on Lemma 1, the proof of Theorem 1 is presented in Appendix A. According to Theorem 1 and Lemma 1, Problem (9) can be equivalently written as

$$\begin{aligned}&\underset{\mathbf{w},\eta}{\mathrm{Max}}\ \ \eta \\ &\text{s.t.}\ \sum_{k=1}^{K} p_k = \sum_{k=1}^{K}\frac{\eta(1+\eta)^{K-k}\sigma^2}{\left|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}\right|^2} \le P, \\ &\quad \|\mathbf{w}\| \le \mathbf{1}, \end{aligned} \tag{15}$$

where $\eta = 2^r - 1$.

Hereto, the first stage to solve Problem (8) is finished, where the optimal power allocation is obtained, and thus the original problem with entangled power allocation and beamforming is reduced to a pure beamforming problem as shown in (15), which will be solved in the next subsection.

B. Beamforming Design With Optimal Power Allocation

The remaining task is to solve Problem (15) and obtain $\mathbf{w}$; then the closed-form expression of $\{p_k\ (k = 1, 2, \cdots, K)\}$ can be obtained by (10). The main challenge is that the first constraint is non-convex, where $\mathbf{w}$ and η are entangled. As the dimension of $\mathbf{w}$, i.e., N, is large in general, it is computationally prohibitive to directly search the optimal solution. However, the introduced variable η is only 1-dimensional. We can search the maximal value of η in the range of $[0, \Gamma]$ with the bisection method, where Γ is the search upper bound. According to the definition of $\eta = 2^r - 1$, η in fact represents the minimal signal to interference plus noise power ratio (SINR) among the K users. If we allocate all the beam gain and power to the user with the best channel condition, i.e., user 1, then user 1 can achieve the highest SINR $\Gamma = (\sum_{n=1}^{N} |[\mathbf{h}_1]_n|)^2 P/(N\sigma^2)$. Thus, we select Γ as the search upper bound. Given a fixed η, we judge whether an appropriate $\mathbf{w}$ can be found in the feasible region of Problem (15). Thus, we need to solve the following problem

$$\begin{aligned}&\underset{\mathbf{w}}{\mathrm{Min}}\ \ f(\mathbf{w}) \triangleq \sum_{k=1}^{K}\frac{\eta(1+\eta)^{K-k}\sigma^2}{\left|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}\right|^2} \\ &\text{s.t.}\ \|\mathbf{w}\| \le \mathbf{1}. \end{aligned} \tag{16}$$

Given η, if the minimal value of the objective function in Problem (16) is no larger than P, which means that a feasible solution can be found with the given η, we enlarge η and solve Problem (16) again. If the minimal value of the objective function in Problem (16) is larger than P, i.e., a feasible solution cannot be found with the given η, we lessen η and solve Problem (16) again. The stopping criterion of the bisection search is that η meets an accuracy requirement.

To solve Problem (16), some approximate manipulations are required to simplify the beamforming problem. Retrospecting the characteristic of the mm Wave channel, the channel response vectors of different users are approximatively orthogonal in general due to the spatial sparsity in the angle domain,[5] which is

$$\frac{\mathbf{h}_m^{\mathrm{H}}}{\|\mathbf{h}_m^{\mathrm{H}}\|}\frac{\mathbf{h}_n}{\|\mathbf{h}_n\|} \approx \begin{cases} 1, & \text{If } m = n; \\ 0, & \text{If } m \ne n. \end{cases} \tag{17}$$

Algorithm 1 AWV Design

Input:
Channel response vectors: $\mathbf{h}_k,\ k = 1, 2, \cdots, K$;
Total transmission power: P;
Noise power: σ^2;
The search accuracy ϵ.
Output:
η and $\mathbf{w}$.

$\eta_{\min} = 0,\ \eta_{\max} = \Gamma$.
while $\eta_{\max} - \eta_{\min} > \epsilon$ **do**
  $\eta = (\eta_{\max} + \eta_{\min})/2$;
  Calculate $\mathbf{w}$ according to (23) and the objective function in Problem (16): $f(\mathbf{w})$.
  if $f(\mathbf{w}) > P$ **then**
    $\eta_{\max} = \eta$.
  else
    $\eta_{\min} = \eta$.
  end if
end while
return η and $\mathbf{w}$.

With this approximation, $\{\frac{\mathbf{h}_k}{\|\mathbf{h}_k\|},\ k = 1, 2, \cdots, K\}$ can be considered as an orthonormal basis of a subspace in $\mathbb{C}^N$. We say the subspace expanded by $\{\frac{\mathbf{h}_k}{\|\mathbf{h}_k\|},\ k = 1, 2, \cdots, K\}$ is a *channel space*. In Problem (16), most beam gains are inclined to focus on the users' directions. Thus, the AWV should be located in the channel space, which can be written as

$$\mathbf{w} = \sum_{k=1}^{K} \alpha_k \frac{\mathbf{h}_k}{\|\mathbf{h}_k\|}, \tag{18}$$

where $\{\alpha_k,\ k = 1, 2, \cdots, K\}$ are the coordinates of $\mathbf{w}$ in the channel space. Substituting (18) into Problem (16), we have

$$\begin{aligned}&\underset{\{\alpha_k\}}{\mathrm{Min}}\ \ \sum_{k=1}^{K}\frac{\eta(1+\eta)^{K-k}\sigma^2}{\alpha_k^2\|\mathbf{h}_k\|^2} \\ &\text{s.t.}\ \sum_{k=1}^{K}\alpha_k^2 = 1. \end{aligned} \tag{19}$$

[5]It is worthwhile noting that if N is large, the probability of two users located in the similar directions is small. Even it happens, the proposed solution is still feasible. It will be shown in the simulations (Fig. 9) that the average achievable rate performance under this assumption is close to the upper bound.

Note that the norm constraint for $\|\mathbf{w}\| \leq \mathbf{1}$ is replaced by $\|\mathbf{w}\| = \mathbf{1}$ here, because the norm of optimal $\mathbf{w}$ is surely 1. Assuming that $\mathbf{w}^\star$ is optimal and $\|\mathbf{w}^\star\| < \mathbf{1}$, we can always normalize the AWV to get a better solution of $\frac{\mathbf{w}^\star}{\|\mathbf{w}^\star\|}$.

To solve Problem (19), we define the Lagrange function as

$$L(\alpha, \lambda) = \sum_{k=1}^{K} \frac{\eta(1+\eta)^{K-k}\sigma^2}{\alpha_k^2 \|\mathbf{h}_k\|^2} + \lambda(\sum_{k=1}^{K} \alpha_k^2 - 1). \quad (20)$$

The Karush-Kuhn-Tucker (KKT) conditions can be obtained by the following equation [37],

$$\begin{cases} \dfrac{\partial L}{\partial \alpha_k} = 0, \quad k = 1, 2, \cdots, K \\ \dfrac{\partial L}{\partial \lambda} = 0. \end{cases} \quad (21)$$

From the KKT conditions, we can obtain the solution of Problem (19), which is given by

$$\begin{aligned} \frac{\partial L}{\partial \alpha_k} = 0 \Rightarrow \frac{-2\eta(1+\eta)^{K-k}\sigma^2}{\alpha_k^3 \|\mathbf{h}_k\|^2} + 2\lambda\alpha_k = 0 \\ \Rightarrow \alpha_k = \sqrt[4]{\frac{\eta(1+\eta)^{K-k}\sigma^2}{\lambda\|\mathbf{h}_k\|^2}} \\ \Rightarrow \alpha_k \propto \sqrt[4]{\frac{\eta(1+\eta)^{K-k}}{\|\mathbf{h}_k\|^2}}. \end{aligned} \quad (22)$$

Thus, the designed AWV in Problem (16) is given by

$$\begin{cases} \bar{\mathbf{w}} = \sum_{k=1}^{K} \sqrt[4]{\dfrac{\eta(1+\eta)^{K-k}}{\|\mathbf{h}_k\|^2}} \dfrac{\mathbf{h}_k}{\|\mathbf{h}_k\|}, \\ \mathbf{w} = \dfrac{\bar{\mathbf{w}}}{\|\bar{\mathbf{w}}\|}. \end{cases} \quad (23)$$

In summary, we give Algorithm 1 to solve Problem (15).

Hereto, we have solved Problem (8) and obtain the solution $\{p_k^\star, \mathbf{w}\}$, where the AWV is obtained in Algorithm 1 and the power allocation is given in (10). The AWV is approximately optimal while the power allocation is optimal for the designed AWV. A leftover problem is to verify the rational of the decoding order. We will consider this problem next.

C. Decoding Order

When formulating Problem (7), we assumed that the decoding order of signals is the increasing order of the channel gains. Next, we will verify that the order of the effective channel gains after beamforming is the same with the channel-gain order. The effective channel gain for user k is

$$\begin{aligned} |\mathbf{h}_k^{\mathrm{H}}\mathbf{w}|^2 \propto |\mathbf{h}_k^{\mathrm{H}}\bar{\mathbf{w}}|^2 &= \left| \sum_{m=1}^{K} \sqrt[4]{\frac{\eta(1+\eta)^{K-m}}{\|\mathbf{h}_m\|^2}} \frac{\mathbf{h}_k^{\mathrm{H}}\mathbf{h}_m}{\|\mathbf{h}_m\|} \right|^2 \\ &\overset{(a)}{=} \left| \sqrt[4]{\frac{\eta(1+\eta)^{K-k}}{\|\mathbf{h}_k\|^2}} \frac{\mathbf{h}_k^{\mathrm{H}}\mathbf{h}_k}{\|\mathbf{h}_k\|} \right|^2 \\ &= \sqrt{\eta(1+\eta)^{K-k}}\|\mathbf{h}_k\|, \end{aligned} \quad (24)$$

where (a) is according to the orthogonal assumption of the channel response vectors. As $\eta = 2^r - 1 > 0$, $\sqrt{\eta(1+\eta)^{K-k}}$ is decreasing for k. We have assumed that the order of the users' channel gains is $\|\mathbf{h}_1\| \geq \|\mathbf{h}_2\| \geq \cdots \geq \|\mathbf{h}_K\|$. Thus, under the orthogonal assumption of the channel response vectors, the order of users' effective channel gains is

$$|\mathbf{h}_1^{\mathrm{H}}\mathbf{w}|^2 \geq |\mathbf{h}_2^{\mathrm{H}}\mathbf{w}|^2 \geq \cdots |\mathbf{h}_K^{\mathrm{H}}\mathbf{w}|^2. \quad (25)$$

As shown in (25), the order of the effective channel gains is the same with that of channel gains. However, this property may not hold if we utilize other decoding orders, which indicates that the increasing-channel-gain decoding order is more reasonable. In the simulations, we will compare the performance of different decoding orders and find that the performance of increasing-channel-gain decoding order is very close to the performance of the optimal decoding order.

D. Consideration of Modulus Constraints

When solving Problem (8), the additional modulus constraints on the AWV were not considered. Next, we will consider the modulus constraints and solve the original problem, i.e., Problem (7). As we have shown in the system model, the modulus constraints on the elements of the AWV are (2) and (3) for SPS and DPS implementations, respectively. Some additional normalized operations on the designed AWV are required to satisfy the constraints. For the SPS implementation, the constant modulus normalization is given by

$$[\mathbf{w}_S]_i = \frac{[\mathbf{w}]_i}{\sqrt{N}\big|[\mathbf{w}]_i\big|}, \ i = 1, 2, \cdots, N. \quad (26)$$

where $\mathbf{w}_S$ denotes the AWV for SPS implementation. For the DPS implementation, the modulus normalization is given by

$$[\mathbf{w}_D]_i = \begin{cases} [\mathbf{w}]_i, & \text{If } \big|[\mathbf{w}]_i\big| \leq \dfrac{2}{\sqrt{N}}; \\ \dfrac{2}{\sqrt{N}} \dfrac{[\mathbf{w}]_i}{\big|[\mathbf{w}]_i\big|}, & \text{If } \big|[\mathbf{w}]_i\big| > \dfrac{2}{\sqrt{N}}. \end{cases} \quad (27)$$

where $\mathbf{w}_D$ denotes the AWV for DPS implementation. Each element of $\mathbf{w}_D$ is the sum weight of the corresponding antenna branch, and it needs to be decomposed into two components, which can be expressed as

$$[\mathbf{w}_D]_i \triangleq a_i e^{j\theta_i} = \frac{1}{\sqrt{N}} e^{j(\theta_i + \varphi_i)} + \frac{1}{\sqrt{N}} e^{j(\theta_i - \varphi_i)}, \quad (28)$$

where $a_i \in [0, \frac{2}{\sqrt{N}}]$ and $\theta_i \in [0, 2\pi)$ are the modulus and the phase of $[\mathbf{w}_D]_i$, respectively, and $\varphi_i = \arccos(\frac{\sqrt{N}a_i}{2})$. Thus, the weights of the two PSs corresponding to $[\mathbf{w}_D]_i$ are

$$\begin{cases} [\tilde{\mathbf{w}}_D]_{2i-1} = \dfrac{1}{\sqrt{N}} e^{j(\theta_i + \varphi_i)}, \\ [\tilde{\mathbf{w}}_D]_{2i} = \dfrac{1}{\sqrt{N}} e^{j(\theta_i - \varphi_i)}. \end{cases} \quad (29)$$

E. Computational Complexity

As we obtained the closed-form optimal power allocation with an arbitrary fixed AWV, the computational complexity is mainly caused by the beamforming algorithm in the second stage. In Algorithm 1, the total search time for η is $T = \log_2(\frac{\Gamma}{\epsilon})$, where Γ is the search upper bound and ϵ is the search accuracy. Thus, the computational complexity of the

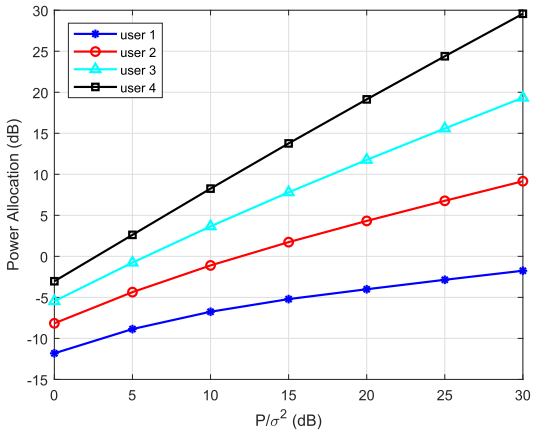

Fig. 2. Power allocation with varying total power to noise ratio, where $N = 32$ and $K = 4$.

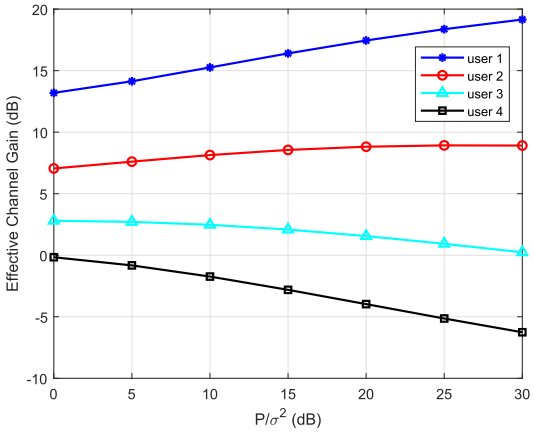

Fig. 3. Effective channel gains with varying total power to noise ratio, where $N = 32$ and $K = 4$.

proposed method is $\mathcal{O}(T)$, which does not increase as N and K. However, if we directly search the solution of Problem (7) and obtain the globally optimal solution, the total complexity is $\mathcal{O}((\frac{1}{\epsilon})^{N+K})$, which exponentially increases as N and K.

IV. Performance Simulations

In this section, we provide simulation results to verify the performance of the proposed joint beamforming and power allocation method in the mm Wave-NOMA system. We adopt the channel model in (4) in the simulations, where the users are uniformly distributed from 10m to 500m away from the BS, and the average channel power, i.e., the mean square value of the complex coefficient, is 1 for the node 100m away from the BS, such that the average channel powers of the other nodes can be computed as their distances to the BS. The number of MPCs for all the users are $L = 4$. Both LOS and NLOS channel models are considered. For the LOS channel, the average power of the NLOS paths is 15 dB weaker than that of the LOS path. For the NLOS channel, the coefficient of each path has an average power of $1/\sqrt{L}$. The search accuracy in Algorithm 1 is $\epsilon = 10^{-6}$.

In the simulations, the minimal achievable rates of "Ideal NOMA/OMA", "SPS-NOMA/SPS-OMA" and "DPS-NOMA/DPS-OMA" are based on the beamforming given in (23), (26) and (27), which are corresponding to the beamforming without CM constraint, with the SPS implementation and with the DPS implementation, respectively. The achievable rate of "Fully digital MIMO" is corresponding to the mm Wave MIMO with zero forcing (ZF) precoding in [16]. The achievable rate of "NOMA upper bound" is corresponding to solving Problem (16) with particle swarm optimization. The density of particles is sufficiently high, and thus the obtained minimal achievable rate can be treated as the upper bound.

We first show the power allocation and the effective channel gains in Figs. 2 and 3, respectively, where the LOS channel model is adopted.[6] Each point is an average result from 10^4 channel realizations. From Fig. 2 we can find that most power is allocated to user 4, the user with the lowest channel gain. Less power is allocated to the users with higher channel gains, so as to reduce interference. Despite all this, it can be observed from Fig. 3 that the effective channel gain of user 4 is still the lowest. The user with a better channel gain have a higher effective channel gain with the proposed solution, which verifies the conclusion in Section III-C about the decoding order. It is noteworthy that the effective channel gains of user 1 and user 4 go increasing and decreasing, respectively, when P/σ^2 becomes higher, which is the result of joint power allocation and beamforming. It indicates that when the total power is high, power and beam gain should be jointly allocated to enlarge the difference of the effective channel gains to achieve a larger minimal user rate.

Next, we compare the performance between the considered mm Wave-NOMA system and a mm Wave-OMA system. We give the following method to calculate the minimal achievable rates in a K-user mm Wave-OMA system, where time division multiple access (TDMA) is used without of generality.

If all the time slots are allocated to user k, the achievable rate for user k is

$$\bar{R}_k = \log_2(1 + \frac{\left|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}\right|^2 P}{\sigma^2}). \tag{30}$$

Assume that the time division is ideal, which means that the time slot can be allocated to the users with any proportion. To maximize the minimal achievable rate of the K users, more time should be allocated to the users with lower channel gains, such that the achievable rates of the K users are equal. Thus, the time allocation for user k is

$$\beta_k = \frac{1/\bar{R}_k}{\sum_{m=1}^{K} 1/\bar{R}_m}. \tag{31}$$

Then the achievable rate of user k in the mm Wave-OMA system is

$$R_k^{\mathrm{OMA}} = \beta_k \bar{R}_k = \frac{1}{\sum_{m=1}^{K} 1/\bar{R}_m}, \tag{32}$$

where all the users have the same achievable rate.

[6]Similar results can be observed when the NLOS channel model is adopted; thus the results are not presented here for conciseness.

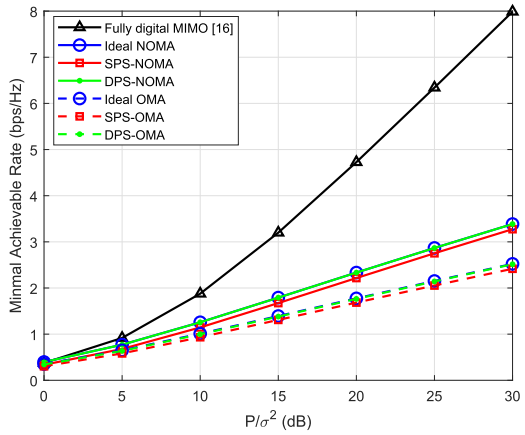

Fig. 4. Comparison of the minimal achievable rates between NOMA, OMA, and MIMO systems with varying total power to noise ratio, where $N = 32$ and $K = 4$.

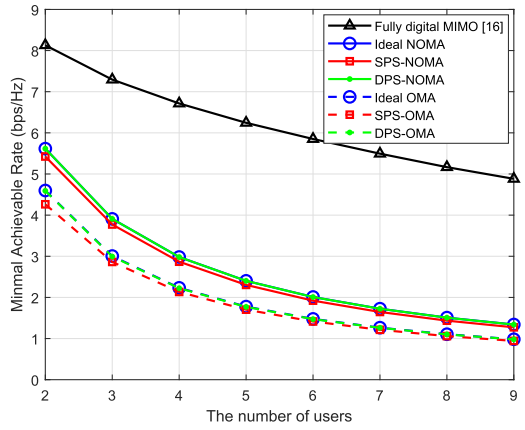

Fig. 5. Comparison of the minimal achievable rates between NOMA, OMA, and MIMO systems with varying number of users, where $N = 32$ and the average transmission power to noise for each user is 20 dB.

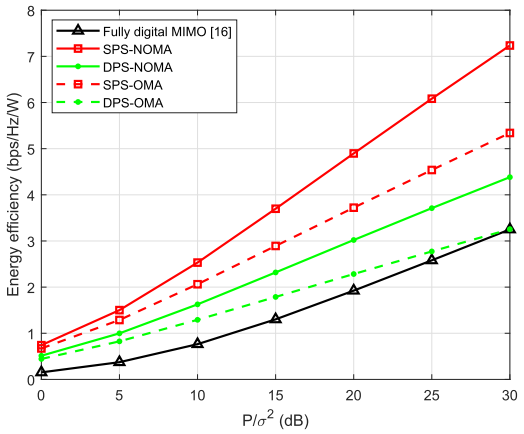

Fig. 6. Comparison of the energy efficiency between NOMA, OMA, and MIMO systems with varying total power to noise ratio, where $N = 32$ and $K = 4$.

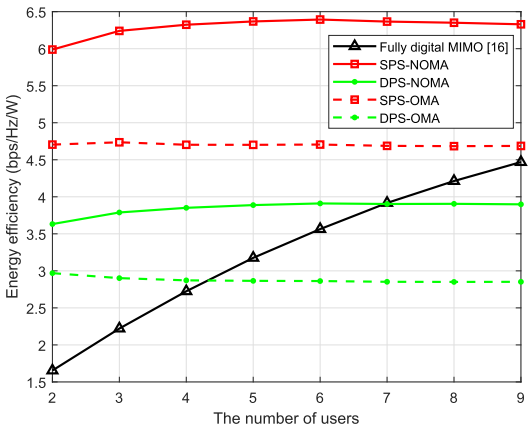

Fig. 7. Comparison of the energy efficiency between NOMA, OMA, and MIMO systems with varying number of users, where $N = 32$ and the average transmission power to noise for each user is 20 dB.

Figs. 4 and 5 show the comparison results of the minimal achievable rates between the mm Wave-NOMA, mm Wave-OMA and mm Wave MIMO systems with varying transmission power to noise ratio and with varying number of users, respectively. Each point in the figures is the average performance of 10^4 LOS channel realizations. We can find that the minimal achievable rates of SPS-NOMA are lower than that of DPS-NOMA, which is very close to the minimal achievable rates of Ideal NOMA. This is because the strict modulus normalization on the AWV for SPS results in significant performance loss, while the modulus normalization on the AWV for DPS is more relaxed and has little impact on the rate performance. In addition, the minimal achievable rates of the mm Wave-NOMA system is distinctly better than those of the mm Wave-OMA system for all the cases, and the superiority is more significant when the total power to noise ratio is higher. In Fig. 4 the rate gain of mm Wave-NOMA over mm Wave-OMA becomes large as the transmission power to noise increases. In Fig. 5, the minimal achievable rates of both mm Wave-NOMA and mm Wave-OMA decrease as the number of users increases. This is mainly due to that the orthogonality of the channel vectors of the users becomes weakened, which deteriorates the beamforming performance and in turn the minimal achievable rate performance. Besides, the minimal achievable rates of the mm Wave-MIMO system are larger than that of the mm Wave-NOMA and mm Wave-OMA systems. The reason is that the number of the RF chain in the proposed mm Wave-NOMA system is 1, while there are N RF chains in the mm Wave MIMO system, which provides a much higher degree of freedom.

Since the circuit power consumptions for the SPS, the DPS, and the fully digital MIMO implementations are quite different, we provide the comparison of energy efficiency in Figs. 6 and 7. The energy efficiency is defined as the radio between the minimum achievable rate and the average power consumption for each user [14], i.e.,

$$\mathrm{EE} = \frac{R_{\min}}{P + N_{\mathrm{RF}} P_{\mathrm{RF}} + N_{\mathrm{PS}} P_{\mathrm{PS}} + P_{\mathrm{BB}}} (\mathrm{bps/Hz/W}), \quad (33)$$

where $P = 30$ mW is the total transmission power, $P_{\mathrm{RF}} = 300$ mW is the power consumed by each RF chain, $P_{\mathrm{PS}} = 40$ mW is the power consumption of each phase shifter, and

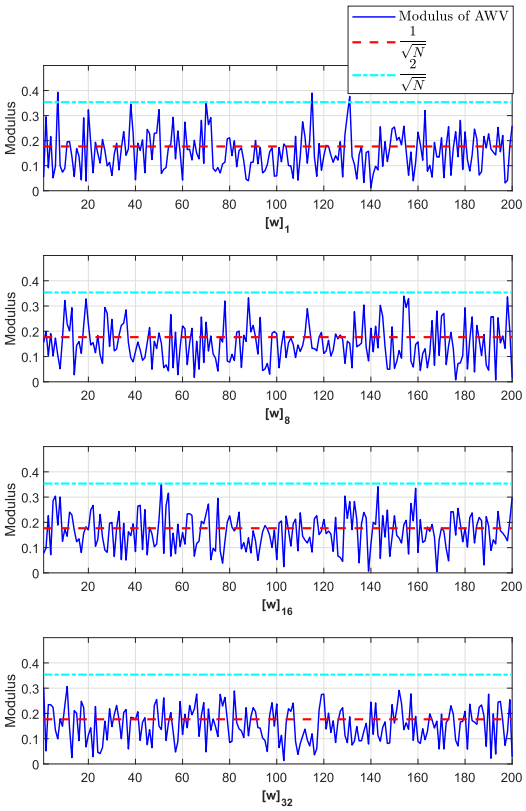

Fig. 8. Modulus of the elements of the AWVs, where $N = 32$, $K = 4$ and $P/\sigma^2 = 25$ dB.

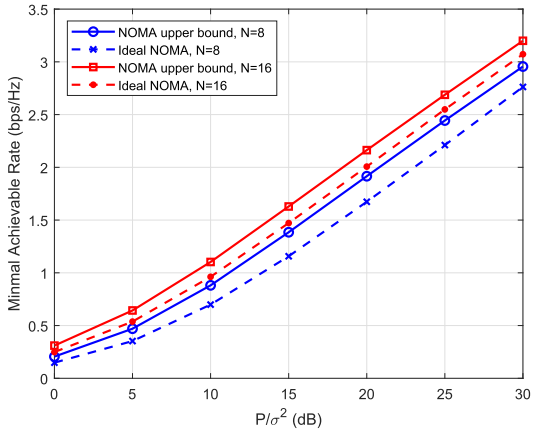

Fig. 9. Comparison of the minimal achievable rates between the proposed solution and the upper bound with varying total power to noise ratio, where $K = 4$.

$P_{\mathrm{BB}} = 200$ mW is the baseband power consumption. N_{RF} is the number of the RF chains, which is equal to 1 for the SPS/DPS implementations and N for the fully digital MIMO implementation. N_{PS} is the number of the PSs, which is equal to N for the SPS implementation and $2N$ for the DPS implementation.

As shown in Fig. 6, the mm Wave-NOMA system can achieve a higher energy efficiency than both the mm Wave-OMA and mm Wave MIMO systems. Since the number of RF chains is equal to the number of antennas at the BS in the mm Wave MIMO system, the high circuit power consumption of the RF chains (300 mW for each one) results in low energy efficiency. On the other hand, the energy efficiency for the SPS implementation is higher than that for the DPS implementation. The reason is that the number of PSs for the DPS implementation is twice more than that for the SPS implementation. Similar results can be obtained in Fig. 7, where the mm Wave-NOMA system has a higher energy efficiency than both the mm Wave-OMA and mm Wave MIMO systems in most instances. However, as the number of users increases, the energy efficiency of mm Wave-NOMA/mm Wave-OMA remains stable, while the energy efficiency of mm Wave MIMO increases. The reason is that the number of RF chains limits the performance of the proposed mm Wave-NOMA system. An alternative approach to increase the spectrum efficiency and the energy efficiency is to use the hybrid beamforming structure in a mm Wave-NOMA system, which will be studied in our future work.

Fig. 8 shows the modulus of the elements of AWVs in (23), where $N = 32$, $K = 4$ and $P/\sigma^2 = 25$ dB. We show the 1st, 8th, 16th and 32th elements of 200 AWVs with different channel realizations. It can be seen that the moduli of the AWV's elements are mainly distributed around $1/\sqrt{N}$, and almost all of them have moduli less than $2/\sqrt{N}$. The reason is as follows. As shown in (23), the solution of AWV is the weighted summation of the normalized channel response vectors of the users. According to the mm Wave channel model in (4), the channel response vector is the weighted summation of the steering vectors, whose elements have unit modulus but different phases. Since $N \gg K$ and $N \gg L_k$, the probability that one element of the AWV has a modulus much larger than the others is low. In other words, all the elements of the AWV may have similar moduli. Thus, after the normalization in (23), the elements of the AWV are distributed around $1/\sqrt{N}$. For this reason, the achievable performance of the DPS-NOMA is close to that of the Ideal NOMA without the modulus constraint, and the modulus normalization for the DPS implementation results in limited performance loss.

In the second stage of the proposed solution, we have assumed that the channel response vectors are orthogonal and then found an appropriate AWV in (16). To evaluate the impact of this approximation, we compare the performance of the proposed solution with the upper-bound performance. Limited by the computational complexity, we provide the simulation results with a relatively small-scale antenna array, i.e., $N = 8, 16$. The comparison result is shown in Fig. 9, where each point is averaged from 10^3 LOS channel realizations. The minimal achievable rate of Ideal NOMA is based on the beamforming given in (23), which is corresponding to the beamforming without the CM constraint and the orthogonality assumption of the channel vectors between the NOMA users. As we can see, when $N = 8$, the performance gap between the proposed solution and the upper bound is no more than 0.25 bps/Hz. When $N = 16$, the performance gap is even smaller, i.e., no more than 0.2 bps/Hz. The reason is that the orthogonality of the channel vectors becomes stronger when N is larger. Thus, the approximation of the beamforming design

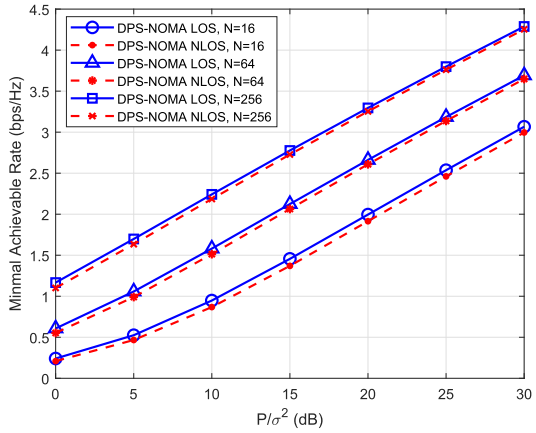

Fig. 10. Performance comparison between LOS and NLOS channel models with varying total power to noise ratio, where $K = 4$.

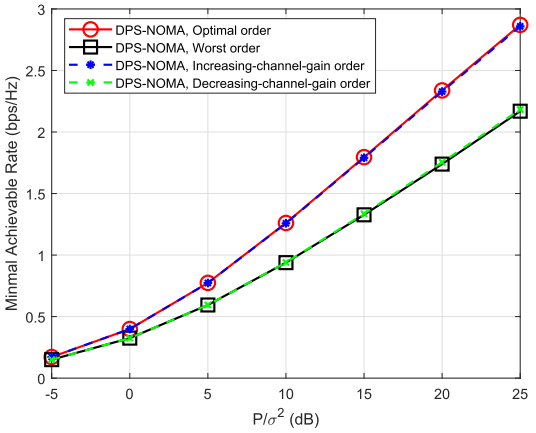

Fig. 11. Comparison of the minimal achievable rates under different decoding orders with varying total power to noise ratio, where $N = 32$ and $K = 4$.

in Problem (16) has limited impact on the system performance, and the proposed sub-optimal solution can achieve an near-upper-bound performance, especially when N is large.

Fig. 10 compares the minimal achievable rates of mm Wave-NOMA under the LOS and NLOS channel models with varying total power to noise ratio. The number of antennas is $N = 16, 64, 256$, respectively. The number of users is $K = 4$. Each point in Fig. 10 is the average performance of 10^4 channel realizations. It can be seen that the performance of DPS-NOMA with the LOS channel model is slightly better than that with the NLOS channel model, because the channel power is more centralized for the LOS channel. However, the performance gap between them is quite small, especially when N is large. The reason is that according to (24), the effective channel gain is linear to $\|\mathbf{h}_k\|$, the norm of the channel vector, rather than that of the power of the strongest path. Thus, the performance gap of DPS-NOMA with the LOS and NLOS channel models is small.

The simulations above are all based on the increasing-channel-gain decoding order. Next, we will show the impact of the decoding order on the mm Wave-NOMA system. Fig. 11 shows the performance comparison between different decoding orders with varying total power to noise ratio, where $N = 32$ and $K = 4$. There are 24 decoding orders in total for the 4 users. Each point in Fig. 11 is the average performance of 10^4 LOS channel realizations. The minimal achievable rates of the 24 decoding orders are all calculated. The order with the highest minimal achievable rate is chosen as the optimal order and the order with the lowest minimal achievable rate is chosen as the worst order. The increasing-channel-gain order is the one adopted in our solution, while the decreasing-channel-gain order is one for comparison. From the figure we can find that there is a significant performance gap between the optimal order and the worst order, which means that the decoding order has an important impact on the performance of mm Wave-NOMA. Moreover, the performance with the increasing-channel-gain order is almost the same as the optimal one, while the performance with the decreasing-channel-gain order is almost the same as the worst one. This result shows the rational of adopting the increasing-channel-gain order in our solution.

V. Conclusion

In this paper, we have investigated downlink max-min fairness mm Wave-NOMA with analog beamforming. A joint beamforming and power allocation problem was formulated and solved in two stages. In the first stage, the closed-form optimal power allocation was obtained with an arbitrary fixed AWV, reducing the joint beamforming and power allocation problem into an equivalent beamforming problem. Then, an appropriate beamforming vector was obtained by utilizing the spatial sparsity in the angle domain of the mm Wave channel. Both implementations of SPS and DPS were considered with different modulus normalizations. The simulation results demonstrate that the modulus normalization has limited impact on the achievable rate performance, especially for the DPS implementation. Moreover, by using the proposed solution, the considered mm Wave-NOMA system can achieve a near-upper-bound performance of the minimal achievable rate, which is significantly better than that of the conventional mm Wave-OMA system.

Appendix A
Proof of Theorem 1

The organization of the proof is as follows. First, given an optimal solution of Problem (9), we can always generate another optimal solution with the expression in (10), which satisfies the condition $R_k = r$ in Theorem 1, and the existence of the optimal solution under the condition $R_k = r$ is proved. Then, we prove the uniqueness of the optimal solution by using the contradiction, where we assume $R_k > r$ and derive a contradiction with the optimality.

Without loss of generality, we denote $\{p_k^\star, r^\star\}$ as one optimal solution of Problem (9) with fixed $\mathbf{w}_0$, where the achievable rate of user k is denoted by $R_k^\star$, and let $\eta^\star = 2^{r^\star} - 1$.

With $\eta^\star$, we can generate another solution $\{p_k^\circ, r^\star\}$ of Problem (9), where

$$\begin{cases} p_1^\circ = \eta^\star \dfrac{\sigma^2}{\left|\mathbf{h}_1^{\mathrm{H}}\mathbf{w}_0\right|^2}, \\ p_2^\circ = \eta^\star (p_1^\circ + \dfrac{\sigma^2}{\left|\mathbf{h}_2^{\mathrm{H}}\mathbf{w}_0\right|^2}), \\ \quad \vdots \\ p_K^\circ = \eta^\star (\sum\limits_{m=1}^{K-1} p_m^\circ + \dfrac{\sigma^2}{\left|\mathbf{h}_K^{\mathrm{H}}\mathbf{w}_0\right|^2}). \end{cases} \tag{34}$$

The following lemma shows that this solution is also an optimal one.

Lemma 2: The solution $\{p_k^\circ, r^\star\}$ is also an optimal solution of Problem (9), and the achievable rates under this parameter setting always satisfy $R_k^\circ = r^\star$ $(1 \le k \le K)$.

Proof: First, we need to verify that the constraints C_0, C_1 and C_2 are all satisfied.

According to the expression of (34), it is obvious that $\{p_k^\circ \ge 0\}$, which means that the constraint C_1 is satisfied.

In addition, according to the assumption that $\{p_k^\star, r^\star\}$ is an optimal solution, we have

$$\begin{aligned} r^\star \le R_k^\star \Rightarrow \eta^\star &\le \frac{\left|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_0\right|^2 p_k^\star}{\left|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_0\right|^2 \sum\limits_{m=1}^{k-1} p_m^\star + \sigma^2} \\ \Rightarrow \eta^\star (\sum_{m=1}^{k-1} p_m^\star &+ \frac{\sigma^2}{\left|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_0\right|^2}) \le p_k^\star. \end{aligned} \tag{35}$$

Next, we use mathematical induction to prove that $p_k^\circ \le p_k^\star$ $(k = 1, 2, \cdots, K)$.

When $k = 1$, according to (35) we have

$$p_1^\circ \le p_1^\star. \tag{36}$$

When $k = n$ $(n \ge 1)$, assume $\{p_1^\circ \le p_1^\star, \cdots, p_n^\circ \le p_n^\star\}$. According to (35) we have

$$\begin{aligned} p_{n+1}^\circ &= \eta^\star (\sum_{m=1}^{n} p_m^\circ + \frac{\sigma^2}{\left|\mathbf{h}_{n+1}^{\mathrm{H}}\mathbf{w}_0\right|^2}) \\ &\le \eta^\star (\sum_{m=1}^{n} p_m^\star + \frac{\sigma^2}{\left|\mathbf{h}_{n+1}^{\mathrm{H}}\mathbf{w}_0\right|^2}) \le p_{n+1}^\star. \end{aligned} \tag{37}$$

Thus, we can conclude that $p_k^\circ \le p_k^\star$ $(k = 1, 2, \cdots, K)$ and we have

$$\sum_{k=1}^{K} p_k^\circ \le \sum_{k=1}^{K} p_k^\star \le P, \tag{38}$$

which means that the constraint C_2 is satisfied.

With the considered solution $(p_k^\circ, r^\star)$, we have

$$\begin{aligned} R_k^\circ &= \log_2(1 + \frac{\left|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_0\right|^2 p_k^\circ}{\left|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_0\right|^2 \sum\limits_{m=1}^{k-1} p_m^\circ + \sigma^2}) \\ &= \log_2(1 + \frac{p_k^\circ}{\sum\limits_{m=1}^{k-1} p_m^\circ + \frac{\sigma^2}{\left|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_0\right|^2}}) \\ &\overset{(a)}{=} \log_2(1 + \eta^\star) = r^\star, \end{aligned} \tag{39}$$

where (a) is based on (34). The above equation means that the constraint C_0 is satisfied.

Since $\{p_k^\circ, r^\star\}$ can satisfy all the constraints, and $R_k^\circ = r^\star$ $(1 \le k \le K)$, it is also an optimal solution of Problem (9). ■

As both $\{p_k^\circ, r^\star\}$ and $\{p_k^\star, r^\star\}$ are optimal solutions of Problem (9), we will prove that they are in fact the same as each other. For this sake, we need to prove that $R_k^\star = r^\star$ $(1 \le k \le K)$. We assume that there exists one user whose achievable is strictly larger than $r^\star$, i.e., $R_{k_0}^\star > r^\star$, and we will prove that this assumption does not hold as follows.

As we have assumed that $R_{k_0}^\star > r^\star$, we have $R_{k_0}^\star > R_{k_0}^\circ = r^\star$. In addition, we have proven that $p_k^\circ \le p_k^\star$ (see the proof in (36) and (37)). According to the expression of R_k in (6), it is straightforward to derive $p_{k_0}^\star > p_{k_0}^\circ$.

We define another solution $\{p_k^\triangle, r^\triangle\}$, where $r^\triangle = r^\star + \delta$, and

$$\begin{cases} p_1^\triangle = \eta^\triangle \dfrac{\sigma^2}{\left|\mathbf{h}_1^{\mathrm{H}}\mathbf{w}_0\right|^2}, \\ p_2^\triangle = \eta^\triangle (p_1^\triangle + \dfrac{\sigma^2}{\left|\mathbf{h}_2^{\mathrm{H}}\mathbf{w}_0\right|^2}), \\ \quad \vdots \\ p_K^\triangle = \eta^\triangle (\sum\limits_{m=1}^{K-1} p_m^\triangle + \dfrac{\sigma^2}{\left|\mathbf{h}_K^{\mathrm{H}}\mathbf{w}_0\right|^2}), \end{cases} \tag{40}$$

where $\eta^\triangle = 2^{r^\triangle} - 1$ and $\delta > 0$. Thus, we have $\eta^\triangle > \eta^\star$.

Next, we prove that $\{p_k^\triangle, r^\triangle\}$ is within the feasible region of Problem (9). Similar to the proof in Lemma 2, we can prove that $\{p_k^\triangle \ge 0\}$ and $R_k^\triangle = r^\triangle > r^\star$ $(1 \le k \le K)$, which means that the constraints C_0 and C_1 are satisfied. According to Lemma 1, the summation of power allocation in (34) and (40) are $g(\eta^\star)$ and $g(\eta^\triangle)$, respectively. As we have proven that $p_{k_0}^\star > p_{k_0}^\circ$, we have $g(\eta^\star) < P$. Otherwise, if $g(\eta^\star) = P$, $\sum\limits_{k=1}^{K} p_k^\star > \sum\limits_{k=1}^{K} p_k^\circ = g(\eta^\star) = P$, which is contradictory to Constraint C_2 in Problem (9). As $g(\eta)$ is an increasing function for η, we can always find a small positive δ, which satisfies $g(\eta^\star + \delta) < P$, i.e., $g(\eta^\triangle) < P$. Thus, the constraint C_2 is satisfied with sufficiently small δ.

In brief, $\{p_k^\triangle, r^\triangle\}$ is within the feasible region of Problem (9) provided that δ is small enough. However, we have $R_k^\triangle = r^\triangle > r^\star$ $(1 \le k \le K)$, which means that the solution $\{p_k^\triangle, r^\triangle\}$ is better than $\{p_k^\star, r^\star\}$, which is contradictory to the fact that $\{p_k^\star, r^\star\}$ is an optimal solution. Thus, the assumption that there exists one user whose achievable is strictly larger than $r^\star$ does not hold. Equivalently, the achievable rates of users under the optimal power allocation satisfy $R_k^\star = r^\star = R_k^\circ$ $(1 \le k \le K)$. Solve the equations set above and we can obtain that $\{p_k^\star, r^\star\}$ is the same as $\{p_k^\circ, r^\star\}$, and the optimal power allocation of Problem (9) is given by (10).

REFERENCES

[1] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[2] V. W. S. Wong, R. Schober, D. W. K. Ng, and L.-C. Wang, *Key Technologies for 5G Wireless Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2017.

[3] T. S. Rappaport *et al.*, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, May 2013.
[4] M. Xiao *et al.*, "Millimeter wave communications for future mobile networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1909–1935, Sep. 2017.
[5] Z. Ding, P. Fan, and H. V. Poor, "On the coexistence of non-orthogonal multiple access and millimeter-wave communications," in *Proc. IEEE Int. Conf. Commun.*, May 2017, pp. 1–6.
[6] A. Benjebbour, Y. Saito, Y. Kishiyama, A. Li, A. Harada, and T. Nakamura, "Concept and practical considerations of non-orthogonal multiple access (NOMA) for future radio access," in *Proc. Int. Symp. Intell. Signal Process. Commun. Syst.*, Nov. 2013, pp. 770–774.
[7] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. IEEE Veh. Technol. Conf.* Dresden, Germany, Jun. 2013, pp. 1–5.
[8] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, Dec. 2014.
[9] L. Dai, B. Wang, Y. Yuan, S. Han, C.-L. I, and Z. Wang, "Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sep. 2015.
[10] Z. Ding, M. Peng, and H. V. Poor, "Cooperative non-orthogonal multiple access in 5G systems," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1462–1465, Aug. 2015.
[11] 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.
[12] L. Zhu, J. Zhang, Z. Xiao, X. Cao, and D. O. Wu, "Optimal user pairing for downlink non-orthogonal multiple access (NOMA)," *IEEE Wireless Commun. Lett.*, vol. 8, no. 2, pp. 328–331, Apr. 2019.
[13] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, "A survey of non-orthogonal multiple access for 5G," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2294–2323, 3rd Quart., 2018.
[14] L. Dai, B. Wang, M. Peng, and S. Chen, "Hybrid precoding-based millimeter-wave massive MIMO-NOMA with simultaneous wireless information and power transfer," *IEEE J. Select. Areas Commun.*, vol. 37, no. 1, pp. 131–141, Jan. 2019.
[15] Z. Ding, P. Fan, and H. V. Poor, "Random beamforming in millimeter-wave NOMA networks," *IEEE Access*, vol. 5, pp. 7667–7681, 2017.
[16] B. Wang, L. Dai, Z. Wang, N. Ge, and S. Zhou, "Spectrum and energy-efficient beamspace MIMO-NOMA for millimeter-wave communications using lens antenna array," *IEEE J. Select. Areas Commun.*, vol. 35, no. 10, pp. 2370–2382, Oct. 2017.
[17] 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, Mar. 2018.
[18] W. Wu and D. Liu, "Non-orthogonal multiple access based hybrid beamforming in 5G mmWave systems," in *Proc. IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun.*, Oct. 2017, pp. 1–7.
[19] D. Zhang, Z. Zhou, C. Xu, Y. Zhang, J. Rodriguez, and T. Sato, "Capacity analysis of NOMA with mmWave massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1606–1618, Jul. 2017.
[20] Z. Xiao, L. Zhu, J. Choi, P. Xia, and X.-G. Xia, "Joint power allocation and beamforming for non-orthogonal multiple access (NOMA) in 5G millimeter wave communications," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2961–2974, May 2018.
[21] L. Zhu, J. Zhang, Z. Xiao, X. Cao, D. O. Wu, and X.-G. Xia, "Joint power control and beamforming for uplink non-orthogonal multiple access in 5G millimeter-wave communications," *IEEE Trans. Wireless Commun.*, vol. 17, no. 9, pp. 6177–6189, Sep. 2018.
[22] Z. Wei, L. Zhao, J. Guo, D. W. K. Ng, and J. Yuan, "Multi-beam NOMA for hybrid mmWave systems," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1705–1719, Feb. 2019.
[23] Z. Xiao, T. He, P. Xia, and X.-G. Xia, "Hierarchical codebook design for beamforming training in millimeter-wave communication," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3380–3392, May 2016.
[24] Z. Xiao, P. Xia, and X.-G. Xia, "Codebook design for millimeter-wave channel estimation with hybrid precoding structure," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 141–153, Jan. 2017.
[25] 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, Oct. 2015.
[26] T. E. Bogale, L. B. Le, A. Haghighat, and L. Vandendorpe, "On the number of RF chains and phase shifters, and scheduling design with hybrid analog-digital beamforming," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3311–3326, May 2016.
[27] Y.-P. Lin, "On the quantization of phase shifters for hybrid precoding systems," *IEEE Trans. Signal Process.*, vol. 65, no. 9, pp. 2237–2246, May 2017.
[28] 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.
[29] J. Umehara, Y. Kishiyama, and K. Higuchi, "Enhancing user fairness in non-orthogonal access with successive interference cancellation for cellular downlink," in *Proc. IEEE Int. Conf. Commun. Syst.*, Nov. 2012, pp. 324–328.
[30] Y. Liu, M. Elkashlan, Z. Ding, and G. K. Karagiannidis, "Fairness of user clustering in MIMO non-orthogonal multiple access systems," *IEEE Commun. Lett.*, vol. 20, no. 7, pp. 1465–1468, Jul. 2016.
[31] H. Xing, Y. Liu, A. Nallanathan, Z. Ding, and H. V. Poor, "Optimal throughput fairness tradeoffs for downlink non-orthogonal multiple access over fading channels," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 3556–3571, Jun. 2018.
[32] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, Jr., "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.
[33] Z. Gao, C. Hu, L. Dai, and Z. Wang, "Channel estimation for millimeter-wave massive MIMO with hybrid precoding over frequency-selective fading channels," *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 1259–1262, Jun. 2016.
[34] Y. Peng, Y. Li, and P. Wang, "An enhanced channel estimation method for millimeter wave systems with massive antenna arrays," *IEEE Commun. Lett.*, vol. 19, no. 9, pp. 1592–1595, Sep. 2015.
[35] P. Wang, Y. Li, L. Song, and B. Vucetic, "Multi-gigabit millimeter wave wireless communications for 5G: From fixed access to cellular networks," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 168–178, Jan. 2015.
[36] J. Lee, G.-T. Gil, and Y. H. Lee, "Exploiting spatial sparsity for estimating channels of hybrid MIMO systems in millimeter wave communications," in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2014, pp. 3326–3331.
[37] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

Zhenyu Xiao (M'11–SM'17) received the B.E. degree from the Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan, China, in 2006, and the Ph.D. degree from the Department of Electronic Engineering, Tsinghua University, Beijing, China, in 2011.

From 2011 to 2013, he held a post-doctoral position at the Electronic Engineering Department, Tsinghua University. From 2013 to 2016, he was a Lecturer at the Department of Electronic and Information Engineering, Beihang University, Beijing, where he is currently an Associate Professor. He has published over 60 papers. His research interests are communication signal processing and practical system implementation for wideband communication systems, which cover synchronization, multipath signal processing, diversity, and multiple antenna technology. He is currently dedicated in millimeter-wave 5G and airborne communications. He has been a TPC member of the IEEE GLOBECOM'12, the IEEE WCSP'12, and the IEEE ICC'15. He has served as a Reviewer for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and the IEEE COMMUNICATIONS LETTERS.

Lipeng Zhu received the B.S. degree from the Department of Mathematics and System Sciences, Beihang University, in 2017, where he is currently pursuing the Ph.D. degree with the Department of Electronic and Information Engineering. His research interests include millimeter-wave communications and non-orthogonal multiple access.

Zhen Gao received the B.S. degree in information engineering from the Beijing Institute of Technology, Beijing, China, in 2011, and the Ph.D. degree in communication and signal processing from the Tsinghua National Laboratory for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, China, in 2016. He is currently an Assistant Professor with the Beijing Institute of Technology. His research interests are in wireless communications, with a focus on multi-carrier modulations, multiple antenna systems, and sparse signal processing.

Dapeng Oliver Wu (S'98–M'04–SM'06–F'13) received the B.E. degree in electrical engineering from the Huazhong University of Science and Technology, Wuhan, China, in 1990, the M.E. degree in electrical engineering from the Beijing University of Posts and Telecommunications, Beijing, China, in 1997, and the Ph.D. degree in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, USA, in 2003.

He is currently a Professor with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, USA. His research interests are in the areas of networking, communications, signal processing, computer vision, machine learning, smart grid, and information and network security.

Dr. Wu was an elected member of the Multimedia Signal Processing Technical Committee and the IEEE Signal Processing Society (2009–2012). He received the University of Florida Term Professorship Award in 2017, the University of Florida Research Foundation Professorship Award in 2009, the AFOSR Young Investigator Program (YIP) Award in 2009, the ONR Young Investigator Program (YIP) Award in 2008, the NSF CAREER award in 2007, the IEEE Circuits and Systems for Video Technology (CSVT) Transactions Best Paper Award for Year 2001, and the Best Paper Awards in the IEEE GLOBECOM 2011 and International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks (QShine) 2006. He has served as the Technical Program Committee (TPC) Chair for the IEEE INFOCOM 2012; and the TPC Chair for the IEEE International Conference on Communications (ICC 2008), the Signal Processing for Communications Symposium, and as a member of the executive committee and/or technical program committee for over 80 conferences. He has served as the Chair for the Award Committee, and the Chair for the Mobile and Wireless Multimedia Interest Group (MobIG), the Technical Committee on Multimedia Communications, and the IEEE Communications Society. He currently serves as the Editor-in-Chief for the IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING, and an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS, and the *IEEE Signal Processing Magazine*. He was the founding Editor-in-Chief of the *Journal of Advances in Multimedia* (2006–2008), and an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He is also the Guest-Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (JSAC), Special Issue on Cross-layer Optimized Wireless Multimedia Communications. He was elected as a Distinguished Lecturer by the IEEE Vehicular Technology Society in 2016.

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

He was a Senior/Research Staff Member at Hughes Research Laboratories, Malibu, CA, USA (1995–1996). In 1996, he joined the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE, USA, 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. He is the author of the book *Modulated Coding for Intersymbol Interference Channels* (New York, Marcel Dekker, 2000).

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 also received the Outstanding Junior Faculty Award of the Engineering School of the University of Delaware in 2001. He is the Technical Program Chair of the Signal Processing Symposium, Globecom 2007 in Washington D.C., USA, and the General Co-Chair of the ICASSP 2005 in Philadelphia, USA. He is currently serving and has served as an Associate Editor for numerous international journals, including the IEEE WIRELESS COMMUNICATIONS LETTERS, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON MOBILE COMPUTING, and the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.