

Random Beamforming in Millimeter-Wave NOMA Networks

Zhiguo Ding, *Senior Member, IEEE*, Pingzhi Fan, *Fellow, IEEE* and H. Vincent Poor, *Fellow, IEEE*

Abstract—This paper investigates the coexistence between two key enabling technologies for fifth generation (5G) mobile networks, non-orthogonal multiple access (NOMA) and millimeter-wave (mmWave) communications. Particularly, the application of random beamforming to mmWave-NOMA systems is considered, in order to avoid the requirement that the base station know all the users' channel state information. Stochastic geometry is used to characterize the performance of the proposed mmWave-NOMA transmission scheme, by using the key features of mmWave systems, e.g., mmWave transmission is highly directional and potential blockages will thin the user distribution. Two random beamforming approaches that can further reduce the system overhead are also proposed, and their performance is studied analytically in terms of sum rates and outage probabilities. Simulation results are also provided to demonstrate the performance of the proposed schemes and verify the accuracy of the developed analytical results.

I. INTRODUCTION

Non-orthogonal multiple access (NOMA) has recently received considerable attention as a promising multiple access (MA) technique to be used in fifth generation (5G) mobile networks [1], [2]. Compared to conventional orthogonal multiple access (OMA), such as time division multiple access and/or frequency division multiple access, NOMA encourages spectrum sharing among multiple users, rather than serving a single user in one orthogonal bandwidth block [3], [4]. Sophisticated power allocation policies and detection methods, such as cognitive radio inspired power allocation, superposition coding and successive interference cancellation (SIC), are used to combat the co-channel interference which is not presented in OMA cases [5], [6]. It is worth pointing out that the use of NOMA can still effectively support massive connectivity and efficiently meet users' diverse QoS requirements, even if the users have similar channel conditions [7].

As an promising enabling technology for 5G networks, NOMA has been shown to be compatible to many other 5G techniques, such as massive multiple-input multiple-output (MIMO), cognitive radio networks, as well as other types of MA techniques, such as orthogonal frequency division multiple access (OFDMA) [8]–[10]. The purpose of this paper is to investigate the coexistence between NOMA and another important 5G technique, millimeter-wave (mmWave) communications [11]–[15]. Even though more bandwidth resources

are available at very high frequencies, the use of NOMA is still important for the following reasons:

- The highly directional feature of mmWave transmission implies that users' channels can be highly correlated, which potentially degrades the system performance. But such correlation is ideal for the application of NOMA.
- The combination supports massive connectivity in dense networks, e.g., where there are hundreds of users to be connected in a small area.
- The rapid growth of mobile Internet services, particularly emerging virtual reality (VR) and augmented reality (AR) services, will dwarf the radio spectrum gains obtained from the mmWave bands, which means that further improvement of the spectral efficiency is still important.

In this paper, we consider a mmWave-NOMA downlink scenario, in which a base station equipped with multiple antennas communicates with multiple single-antenna nodes. While MIMO-NOMA has been extensively studied in [16]–[18], the application of mmWave communications makes the addressed MIMO-NOMA scenario much different, mainly due to the characteristics of mmWave propagation. The contributions of this paper are four-fold:

- We first consider the application of random beamforming to the addressed mmWave-NOMA scenario, in which a single beam is randomly generated by the base station. While random beamforming does not require the base station to know all the users' channel vectors, conventional random beamforming still requires all the users to send their scale channel gains to the base station, which can consume significant system overhead in a network with a large number of users. The fact that mmWave transmission is highly directional is used in this paper to avoid scheduling those users who are likely to have low signal strength, which reduces the number of users who need to feed their channel quality information back to the base station and hence reduces the system overhead. Stochastic geometry is applied to characterize the sum rate and the outage probabilities achieved by the proposed beamforming scheme, where the blockage feature of mmWave propagation is also used to model the user distribution more realistically.
- In a fast time varying situation, in which the phases and the amplitudes of the users' channel gains change rapidly, a low-feedback transmission scheme is proposed by assuming that only the users' distance information is available to the base station. As a result, the users are ordered according to their path losses, instead of their effective channel gains. The impact of this partial channel state information (CSI) on the performance of the mmWave-NOMA downlink network is investigated.
- A one-bit feedback random beamforming scheme is also proposed in order to further reduce the system overhead.

Parts of the material in this paper were presented at IEEE International Conference in Communications, France, France. Z. Ding and H. V. Poor are with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA. Z. Ding is also with the School of Computing and Communications, Lancaster University, LA1 4WA, UK. P. Fan is with the Institute of Mobile Communications, Southwest Jiaotong University, Chengdu, China.

The work of Z. Ding was supported by the UK Engineering and Physical Science Research Council under grant number EP/L025272/1 and by the European Commission Program H2020-MSCA-RISE-2015 under grant number 690750. The work of P. Fan was supported by the Huawei HIRP Flagship Project (No.YB201504), the National Science and Technology Major Project (No.2016ZX03001018-002) and the 111 project (No.111-2-14). The work of H. V. Poor was supported by the U.S. National Science Foundation under Grant CCF-1420575 and CNS-1456793.

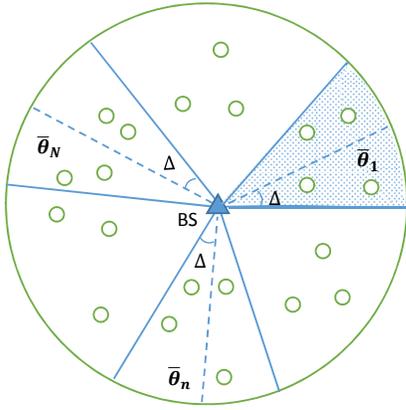


Fig. 1. A system diagram for the addressed mmWave-NOMA scenario. $\bar{\theta}_n$ denotes a beamforming vector randomly generated by the base station. Only the users that fall into a specific wedge-shaped sector will be scheduled, which is to ensure that the maximal angle difference between a scheduled user's channel vector and its associated beam is Δ .

In particular, the base station sets a threshold which is broadcast to the users. Each user feeds one bit back to the base station to indicate its channel quality. The use of one-bit feedback can effectively reduce the amount of feedback, but will cause an ordering ambiguity at the base station. The impact of this ambiguity on the performance of the one-bit feedback transmission scheme is investigated. Furthermore, the effect of the threshold is also characterized, where the obtained analytical results show that a properly designed threshold can ensure that the full diversity gain is achievable by the user selected to be the NOMA strong user.

- The performance for the more challenging scenario in which the base station generates multiple orthonormal beams is also investigated. Compared to the case with a single beam, each user in the scenario with multiple beams suffers more interference, including intra NOMA group interference and inter-beam interference. Because mmWave transmission is highly directional, inter-beam interference can be effectively suppressed by scheduling the users whose channel vectors are aligned with the randomly generated beams. Exact expressions for the outage probabilities achieved by the random beamforming scheme and their approximations are developed in order to obtain greater insights.

II. SYSTEM MODEL

Consider a mmWave-NOMA downlink transmission scenario with one base station communicating with multiple users, as shown in Fig. 1. The base station is equipped with M antennas and each user has a single antenna. Denote the disk covered by the base station by \mathcal{D} . Assume that the base station is located at the origin of \mathcal{D} and denote the radius of the disk by $R_{\mathcal{D}}$. Assume that users are randomly deployed in the disc following a homogeneous Poisson point process (HPPP) with density λ [19]. Therefore, the number of users in the disk is Poisson distributed, i.e., $P(K \text{ users in } \mathcal{D}) = \frac{\mu^K e^{-\mu}}{K!}$, where $\mu = \pi R_{\mathcal{D}}^2 \lambda$.

As discussed in [12] and [13], the mmWave channel model is quite different from those of conventional lower frequency

cellular networks; in particular, the mmWave-based channel vector from the base station to user k can be expressed as follows:

$$\mathbf{h}_k = \sqrt{M} \frac{a_{k,0} \mathbf{a}(\theta_k^0)}{\sqrt{1 + d_k^{\alpha_{LOS}}}} + \sqrt{M} \sum_{l=1}^L \frac{a_{k,l} \mathbf{a}(\theta_k^l)}{\sqrt{1 + d_k^{\alpha_{NLOS}}}}, \quad (1)$$

where L is the number of multi-paths,

$$\mathbf{a}(\theta) = \frac{1}{\sqrt{M}} [1 \quad e^{-j\pi\theta} \quad \dots \quad e^{-j\pi(M-1)\theta}]^T, \quad (2)$$

d_k denotes the distance between the transceivers, α_{NLOS} and α_{LOS} denote the path loss exponents for the non-line-of-sight (NLOS) and line-of-sight (LOS) paths, respectively, $a_{k,l}$ denotes the complex gain for the l -th path and is complex Gaussian distributed, i.e., $a_{k,l} \sim CN(0,1)$, and θ_k^l denotes the normalized direction of the l -th path. We assume that the channel gains are independent from path to path. For notational simplicity, the normalized direction of a path is treated the same as its physical angle of departure, and in Section III-A, we can show this simplification has no impact on the performance of the proposed mmWave-NOMA scheme.

As discussed in [13] and [20], in mmWave communications, the effect of LOS links is dominant, compared to those of NLOS links, e.g., the gain of an LOS link can be 20 dB stronger than those of NLOS links. Therefore the first factor at the right-hand side of (1) is dominant, which yields the following simplified channel model:

$$\mathbf{h}_k = \sqrt{M} \frac{a_k \mathbf{a}(\theta_k)}{\sqrt{1 + d_k^{\alpha}}}, \quad (3)$$

where the subscripts 0 and LOS have been omitted to simplify the notation.

In practice, the direct path between the mmWave transceivers might be blocked by obstacles, which means that an LOS path does not always exist. As a result, in addition to path loss and fading attenuation, mmWave transmission also suffers potential blockages, which is an important feature to be captured. A simple way to model these blockages is to assume the existence of an LOS path if the distance between the transceivers is smaller than a threshold [14] and [21]. Alternatively, a more sophisticated way to model the probability of there being an LOS path for mm-Wave transmission has been introduced in [15], [22] as follows:

$$P(LOS) = e^{-\phi d_k}, \quad (4)$$

where ϕ is determined by the building density, the shape of the buildings, etc. In this paper, we will use (4) for modelling blockages in the addressed mmWave communication scenario. It is important to point out that these blockages will thin the node distribution, which will be discussed in detail in the next section.

III. RANDOM BEAMFORMING: A SINGLE-BEAM CASE

Many existing precoding and beamforming schemes for NOMA require that the base station has access to the users' CSI. These approaches can consume a substantial amount of system overhead, if there are many users in the system. In order to reduce the system overhead, we consider the application

of random beamforming to mmWave-NOMA communication scenarios.

A. The Application of Random Beamforming to NOMA

In this section, we focus on the case in which a single beam, denoted by \mathbf{p} , is generated at the base station. Note that in the context of mmWave communications, analog precoding is preferable compared to digital precoding since the amplitude of a signal is kept constant and only the phase is changed. Therefore, following [13] and [23], we use the following choice for beamforming:

$$\mathbf{p} = \mathbf{a}(\bar{\theta}), \quad (5)$$

where $\bar{\theta}$ is uniformly distributed between -1 and 1 . This choice of precoding is analog precoding since it alters the signal phase only, and keeps the signal modulus constant. It is worth pointing out that this beamformer is also a special case of the hybrid precoding design in [24] with one radio frequency chain and M antennas.

One straightforward solution for user scheduling is to ask each user to feed its effective channel gain $|\mathbf{h}_j^H \mathbf{p}|^2$ back to the base station, and then the base station schedules the user with the strongest channel. However, such an approach will still consume considerable system overhead, particularly if there are many users in the cell.

In the context of mmWave communications, a useful observation is that many users do not have to participate in the competition for access to the channel, as explained in the following. Without loss of generality, user j is randomly chosen to be served on beam \mathbf{p} . The effective channel gain of this user on the randomly generated beam, $|\mathbf{h}_j^H \mathbf{p}|^2$, can be written as follows:

$$|\mathbf{h}_j^H \mathbf{p}|^2 = M \frac{|a_j|^2 |\mathbf{p}^H \mathbf{a}(\theta_j)|^2}{1 + d_j^\alpha} = \frac{|a_j|^2 \left| \sum_{l=0}^{M-1} e^{-j\pi l(\bar{\theta} - \theta_j)} \right|^2}{M(1 + d_j^\alpha)}. \quad (6)$$

Following steps similar to those in [13], this effective channel gain can be rewritten as follows:

$$\begin{aligned} |\mathbf{h}_j^H \mathbf{p}|^2 &= \frac{|a_j|^2 \sin^2 \left(\frac{\pi M(\bar{\theta} - \theta_j)}{2} \right)}{M(1 + d_j^\alpha) \sin^2 \left(\frac{\pi(\bar{\theta} - \theta_j)}{2} \right)} \\ &= \frac{|a_j|^2}{(1 + d_j^\alpha)} F_M(\pi[\bar{\theta} - \theta_j]), \end{aligned} \quad (7)$$

where $F_M(x)$ denotes the Fejér kernel. Note that the Fejér kernel goes to zero quickly by increasing its argument, i.e., $F_M(x) \rightarrow 0$ for increasing x . This means that a user can have a large effective channel gain on beam \mathbf{p} if this user's channel vector is aligned with the direction of the beam.

Following this observation, we will schedule only the users who are located in the wedge-shaped sector served by the beam, as highlighted in Fig. 1. Particularly, this sector is denoted by \mathcal{D}_θ , and its central angle is 2Δ , which means that the maximal angle difference between a scheduled user's channel vector and the beam is Δ , and $\Delta \rightarrow 0$ is required to ensure a large effective channel gain. Note that, when

$\Delta \rightarrow 0$, the use of the normalized direction of a path to replace its physical angle of departure has no impact on the performance of the proposed scheme, as illustrated in the following. Recall that the normalized direction θ is a function of the physical angle of departure, denoted by ϕ_θ , i.e., $\theta = \frac{2d \sin(\phi_\theta)}{\lambda}$, where λ and d are the carrier wavelength and the antenna separation distance, respectively. If $\Delta \rightarrow 0$, we have $|\bar{\theta} - \theta_j| \rightarrow 0$, and hence $|\phi_{\bar{\theta}} - \phi_{\theta_j}| \rightarrow 0$, which means that the two physical angles are very similar if the two normalized directions are similar. Furthermore, as $\Delta \rightarrow 0$, the application of Taylor series leads to $\theta_j - \bar{\theta} \approx \frac{2d \cos(\phi_{\bar{\theta}})}{\lambda} (\phi_{\theta_j} - \phi_{\bar{\theta}})$, and so our analytical results based on the normalized directions can be extended to the case with the physical angles in a straightforward manner.

B. The Implementation of NOMA

Suppose that there are K users in the sector, \mathcal{D}_θ , and these users are ordered according to their effective channel gains as follows:

$$|\mathbf{h}_1^H \mathbf{p}|^2 \leq \dots \leq |\mathbf{h}_K^H \mathbf{p}|^2. \quad (8)$$

Similarly to [6] and [16], we consider the case in which two users will be selected for the implementation of NOMA. Note that the implementation of NOMA in long term evolution advanced (LTE-A) is also based on the two-user case [25]. Since the aim of this paper is to study the impact of NOMA on mmWave communications, without loss of generality, we assume that user i and user j are paired together for NOMA transmission on a randomly generated beam. Note that i and j can be chosen arbitrarily, constrained by $1 \leq i < j \leq K$. As a result, the performance of mmWave-NOMA with different scheduled users can be investigated, and the insights obtained from the performance analysis can offer guidelines for the design of practical user scheduling algorithms. Therefore, the signal sent by the base station is given by

$$\mathbf{p} (\beta_i s_i + \beta_j s_j), \quad (9)$$

where β_i denotes the power allocation coefficient. Since $|\mathbf{h}_i^H \mathbf{p}|^2 < |\mathbf{h}_j^H \mathbf{p}|^2$, the application of NOMA means $\beta_i \geq \beta_j$, where $\beta_i^2 + \beta_j^2 = 1$.

Therefore, user i will receive the following observation:

$$y_i = \mathbf{h}_i^H \mathbf{p} (\beta_i s_i + \beta_j s_j) + n_i, \quad (10)$$

where n_i denotes additive Gaussian noise. User i will treat its partner's message as noise and directly decode its information with the following signal-to-interference-plus-noise ratio (SINR):

$$\text{SINR}_i = \frac{|\mathbf{h}_i^H \mathbf{p}|^2 \beta_i^2}{|\mathbf{h}_i^H \mathbf{p}|^2 \beta_j^2 + \frac{1}{\rho}}, \quad (11)$$

where ρ denotes the transmit signal-to-noise ratio (SNR). As a result, the outage probability for user i to decode its information is given by

$$P_{i|K}^o = \text{P}(\log(1 + \text{SINR}_i) < R_i | K) = \text{P}(\text{SINR}_i < \epsilon_i | K), \quad (12)$$

which is conditioned on the number of users in \mathcal{D}_θ , where $\epsilon_i = 2^{R_i} - 1$.

User j first tries to decode its partner's message with the following SINR: $\text{SINR}_{i \rightarrow j} = \frac{|\mathbf{h}_j^H \mathbf{p}|^2 \beta_i^2}{|\mathbf{h}_j^H \mathbf{p}|^2 \beta_j^2 + \frac{1}{\rho}}$. If $\text{SINR}_{i \rightarrow j} \geq \epsilon_i$, the user can decode its own message with the following SNR:

$$\text{SINR}_j = \rho |\mathbf{h}_j^H \mathbf{p}|^2 \beta_j^2, \quad (13)$$

after removing its partner's information, a procedure known as SIC. Therefore the outage probability experienced by user j can be expressed as follows:

$$P_{j|K}^o = 1 - \text{P}(\text{SINR}_{i \rightarrow j} > \epsilon_i, \text{SINR}_j > \epsilon_j | K), \quad (14)$$

which is again conditioned on K .

As a result, the outage sum rate achieved by the mmWave-NOMA transmission scheme can be expressed as follows:

$$R_{sum}^{NOMA} = \text{P}(K=1)(1 - P_{OMA}^{1|K})R_1 + \sum_{k=2}^{\infty} \text{P}(K=k) \times \left((1 - P_{i|K}^o)R_i + (1 - P_{j|K}^o)R_j \right), \quad (15)$$

and the sum rate achieved by mmWave-OMA can be expressed similarly as follows:

$$R_{sum}^{OMA} = \text{P}(K=1)(1 - P_{OMA}^{1|K})R_1 + \sum_{k=2}^{\infty} \text{P}(K=k) \times \left((1 - P_{OMA}^{i|K})R_i + (1 - P_{OMA}^{j|K})R_j \right), \quad (16)$$

where $P_{OMA}^{n|K}$ denotes the conditional outage probability when OMA is used. The reason for using the OMA mode in (15) is that it is possible to have a single user in \mathcal{D}_θ . In this case, NOMA cannot be implemented and we simply use OMA, i.e., $P_{OMA}^{n|K} = \text{P}(\log(1 + \rho |\mathbf{h}_n^H \mathbf{p}|^2) < 2R_n)$, for $n \in \{i, j\}^1$.

C. Characterization of the Sum Rate and Outage Probabilities

In order to evaluate the sum rate shown in (15), it is important to find expressions for the outage probabilities, $P_{j|K}^o$ and $P_{i|K}^o$; these are related to the probability density function (pdf) of the ordered channel gain, $|\mathbf{h}_j^H \mathbf{p}|^2$, which is provided in the following lemma.

Lemma 1. *Suppose that there are K users in \mathcal{D}_θ . The pdf of the ordered channel gain, $|\mathbf{h}_j^H \mathbf{p}|^2$, is given by*

$$f_{|\mathbf{h}_j^H \mathbf{p}|^2}(z) = c_j \frac{dF_{\pi(j)}(z)}{dz} F_{\pi(j)}^{j-1}(z) (1 - F_{\pi(j)}(z))^{K-j}, \quad (17)$$

where $c_j = \frac{K!}{(j-1)!(K-j)!}$,

$$F_{\pi(j)}(y) = \int_{\bar{\theta}-\Delta}^{\bar{\theta}+\Delta} \int_0^{R_{\mathcal{D}}} \left(1 - e^{-\frac{y(1+r(x)^\alpha)}{F_M(\pi[\bar{\theta}-\theta_{\pi(j)}])}} \right) \times \frac{\lambda \phi^2 e^{-\phi r}}{2\Delta \lambda \gamma(2, R_{\mathcal{D}} \phi)} r dr d\theta, \quad (18)$$

¹One can also use $P_{OMA}^{1|K} = \text{P}(\log(1 + \rho |\mathbf{h}_1^H \mathbf{p}|^2) < R_1)$ for the case $K=1$, which will make the notation in (15) and (16) more complicated. It is worth pointing out that the probability of having $K=1$ is very small and different designs for this trivial case do not cause much difference to the overall sum rate.

and $\gamma(\cdot)$ denotes the incomplete gamma function.

Proof: The density function in the lemma can be evaluated by first characterizing the unordered channel gains and then applying the theory of order statistics.

First we focus on an unordered channel gain, denoted by $|\mathbf{h}_{\pi(j)}^H \mathbf{p}|^2$. Denote the location of this node by $x_{\pi(j)}$, where its probability distribution and pdf are denoted by $P_{X_{\pi(j)}}$ and $p_{X_{\pi(j)}}$, respectively. In this case we can find the cumulative distribution function (CDF) of the unordered channel gain as follows:

$$F_{\pi(j)}(y) = \int_{\mathcal{D}_\theta} \text{P}\left(|\mathbf{h}_{\pi(j)}^H \mathbf{p}|^2 < y \mid X_{\pi(j)} = x_{\pi(j)}\right) dP_{X_{\pi(j)}} \\ = \int_{\mathcal{D}_\theta} \left(1 - e^{-\frac{y(1+r(x)^\alpha)}{F_M(\pi[\bar{\theta}-\theta_{\pi(j)}])}} \right) p_{X_{\pi(j)}}(x) dx,$$

where $r(x)$ denotes the distance from the origin to point x . Note that the conditioning on K has been omitted since it does not affect the CDF.

It is important to note that the nodes participating in NOMA no longer follow the original HPPP with parameter λ , because of potential blockages. Particularly, with the blockage model in (4), it is less likely for a user far away from the base station to have an LOS path. Therefore, following the discussions in [26], the effect of blockages is to thin the original homogeneous point process and this thinning process yields another PPP with the following intensity:

$$\lambda_{\Phi_2}(x) = \lambda e^{-\phi r(x)}. \quad (19)$$

Therefore, the mean measure for this new PPP, denoted by $\mu_{\Phi_2}(\mathcal{D}_\theta)$, can be obtained as follows:

$$\mu_{\Phi_2}(\mathcal{D}_\theta) = \int_{\mathcal{D}_\theta} \lambda_{\Phi_2}(x) dx \\ = \int_{\bar{\theta}-\Delta}^{\bar{\theta}+\Delta} \int_0^{R_{\mathcal{D}}} \lambda e^{-\phi r} r dr d\theta = 2\Delta \lambda \phi^{-2} \gamma(2, R_{\mathcal{D}} \phi). \quad (20)$$

As a result, after considering potential blockages, the probability of having K users in the sector, \mathcal{D}_θ , can be obtained as follows:

$$\text{P}(K=k) = \frac{(\mu_{\Phi_2}(\mathcal{D}_\theta))^k}{k!} e^{-\mu_{\Phi_2}(\mathcal{D}_\theta)}. \quad (21)$$

Since the intensity and the mean measure of the new PPP are known, the pdf of $x_{\pi(j)}$ can be written as follows:

$$p_{X_{\pi(j)}}(x) = \frac{\lambda_{\Phi_2}(x)}{\mu_{\Phi_2}(\mathcal{D}_\theta)} = \frac{\lambda \phi^2 e^{-\phi r(x)}}{2\Delta \lambda \gamma(2, R_{\mathcal{D}} \phi)}. \quad (22)$$

Accordingly, the CDF of the unordered channel gain can be written as follows:

$$F_{\pi(j)}(y) = \int_{\mathcal{D}_\theta} \left(1 - e^{-\frac{y(1+r(x)^\alpha)}{F_M(\pi[\bar{\theta}-\theta_{\pi(j)}])}} \right) \frac{\lambda \phi^2 e^{-\phi r(x)}}{2\Delta \lambda \gamma(2, R_{\mathcal{D}} \phi)} dx,$$

and by using polar coordinates, the expression for $F_{\pi(j)}(y)$ in the lemma can be obtained. After using the assumption that all the channel gains are independent and identically distributed

and also applying the theory of order statistics [27], the proof is complete. ■

By applying the above lemma and also some algebraic manipulations, $P_{j|K}^o$ and $P_{i|K}^o$ can be obtained in the following corollary.

Corollary 1. *By using the proposed mmWave-NOMA transmission scheme, the outage probability experienced by user j conditioned on K is given by*

$$P_{j|K}^o = c_j \sum_{p=0}^{K-j} \binom{K-j}{p} (-1)^p \frac{F_{\pi(j)}^{j+p}(\eta_j)}{j+p}, \quad (23)$$

if $\beta_i^2 > \beta_j^2 \epsilon_i$, otherwise $P_{j|K}^o = 1$, where $\eta_j = \max \left\{ \frac{\epsilon_i}{\beta_i^2 - \beta_j^2 \epsilon_i}, \frac{\epsilon_j}{\rho \beta_j^2} \right\}$. The conditional outage probability for user i is given by

$$P_{i|K}^o = c_i \sum_{p=0}^{K-i} \binom{K-i}{p} (-1)^p \frac{F_{\pi(j)}^{i+p}(\eta_i)}{i+p}, \quad (24)$$

where $\eta_i = \frac{\epsilon_i}{\beta_i^2 - \beta_j^2 \epsilon_i}$.

By using the above corollary and substituting (21), (23) and (24) into (15) and (16), the sum rates achieved by mmWave-NOMA and mmWave-OMA can be calculated.

D. Asymptotic Performance Analysis

The obtained results shown in (23) and (24) are quite complicated, since they involve the calculation of double integrals. In order to obtain some insight, we will obtain approximations to these expressions. Particularly, our asymptotic studies are carried out by using the following two assumptions. One is that the central angle of the sector, 2Δ , is small, i.e., $\Delta \rightarrow 0$, and the other is the high SNR assumption. The use of these two assumptions leads to the following lemma.

Lemma 2. *When $\Delta \rightarrow 0$ and at high SNR, the conditional outage probabilities $P_{i|K}^o$ and $P_{j|K}^o$ can be approximated as follows:*

$$P_{k|K}^o \approx c_k \frac{F_{\pi(j)}^{k+p}(\eta_k)}{k}, \quad (25)$$

where $k \in \{i, j\}$. The diversity gain available at user k is k .

Proof: In order to use the assumption $\Delta \rightarrow 0$, recall that the Fejér kernel can be written as follows:

$$F_M(\pi[\bar{\theta} - \theta]) = \frac{\sin^2\left(\frac{\pi M(\bar{\theta} - \theta)}{2}\right)}{M \sin^2\left(\frac{\pi(\bar{\theta} - \theta)}{2}\right)}. \quad (26)$$

Note that $|\bar{\theta} - \theta| \leq \Delta$. When Δ is small, the Fejér kernel can be approximated as follows:

$$\begin{aligned} F_M(\pi[\bar{\theta} - \theta]) &\approx M \operatorname{sinc}^2\left(\frac{\pi M(\bar{\theta} - \theta)}{2}\right) \\ &\approx M \left(1 - \frac{\pi^2 M^2 (\bar{\theta} - \theta)^2}{12}\right), \end{aligned} \quad (27)$$

where the first approximation follows from $\sin(x) \approx x$ for $x \rightarrow 0$, and the second approximation is due to the two following facts: $\operatorname{sinc}(x) \approx 1 - \frac{x^2}{6}$ and $(1-x)^2 \approx 1-2x$, for $x \rightarrow 0$ [28].

Therefore, the CDF of an unordered channel gain can be approximated as follows:

$$\begin{aligned} F_{\pi(j)}(y) &\approx \int_{\bar{\theta}-\Delta}^{\bar{\theta}+\Delta} \int_0^{R_D} \frac{\lambda \phi^2 e^{-\phi r}}{2\Delta \lambda \gamma(2, R_D \phi)} \\ &\times \left(1 - e^{-\frac{y(1+r^\alpha)}{M \left(1 - \frac{\pi^2 M^2 (\bar{\theta} - \theta)^2}{12}\right)}}\right) r dr d\theta \\ &\approx \int_{\bar{\theta}-\Delta}^{\bar{\theta}+\Delta} \int_0^{R_D} \frac{\lambda \phi^2 e^{-\phi r}}{2\Delta \lambda \gamma(2, R_D \phi)} \\ &\times \left(1 - e^{-\frac{y(1+r^\alpha)}{M} \left(1 + \frac{\pi^2 M^2 (\bar{\theta} - \theta)^2}{12}\right)}\right) r dr d\theta, \end{aligned} \quad (28)$$

where the last approximation follows from $(1-x)^{-1} \approx 1+x$, for $x \rightarrow 0$.

After applying the assumption that $\Delta \rightarrow 0$, we will further apply the high SNR approximation. Note that at high SNR, both η_i and η_j go to zero, which means

$$\begin{aligned} F_{\pi(j)}(\eta_i) &\approx \int_{\bar{\theta}-\Delta}^{\bar{\theta}+\Delta} \int_0^{R_D} \frac{\lambda \phi^2 e^{-\phi r}}{2\Delta \lambda \gamma(2, R_D \phi)} \left(\frac{\eta_i(1+r^\alpha)}{M}\right) \\ &\times \left(1 + \frac{\pi^2 M^2 (\bar{\theta} - \theta)^2}{12}\right) r dr d\theta. \end{aligned} \quad (29)$$

After some algebraic manipulations, the CDF for an unordered channel gain can be approximated as follows:

$$\begin{aligned} F_{\pi(j)}(\eta_i) &\approx \frac{\eta_i}{2M \gamma(2, R_D \phi)} \left(2 + \frac{\pi^2 M^2 \Delta^2}{18}\right) \\ &\times (\gamma(2, R_D \phi) + \phi^{-\alpha} \gamma(\alpha+2, R_D \phi)) \sim \frac{1}{\rho}, \end{aligned} \quad (30)$$

where $f(\rho) \sim \frac{1}{\rho^x}$ when $\lim_{\rho \rightarrow \infty} \frac{\log(f(\rho))}{\log \rho} = -x$ [29].

By using the above approximations, the outage probability at user i can be approximated as follows:

$$\begin{aligned} P_{i|K}^o &= c_i \sum_{p=0}^{K-i} \binom{K-i}{p} (-1)^p \frac{F_{\pi(j)}^{i+p}(\eta_i)}{i+p} \\ &\approx c_i \frac{F_{\pi(j)}^{i+p}(\eta_i)}{i} \sim \frac{1}{\rho^i}, \end{aligned} \quad (31)$$

which means that the diversity gain at user i is i . The results for user j can be obtained similarly, and the proof is complete. ■

Remark: Note that an implication of having a small Δ is that the area of the sector becomes so small that there might be no user in it. But in many practical scenarios, such as in a sport stadium or a conference hall, the users are so densely deployed that it is always possible to find multiple users located in a sector even with a small Δ .

IV. RANDOM BEAMFORMING WITH LIMITED FEEDBACK

In the previous section, it is assumed that the base station has perfect knowledge of the users' effective channel gains. However, for a fast time varying situation, this assumption might not be realistic, since the phases of the channel vectors and their fading coefficients, θ_k and a_k , are changing rapidly. In this section, we investigate two random beamforming schemes with low system overhead.

A. With the Distance Information Available at the Base Station

Compared to the phases and fading coefficients of the channels, the users' distance information will change relatively slowly, which means that it is more realistic for the base station to have access to the users' distance information only. Therefore, in this subsection, we investigate the impact of this partial CSI on the performance of mmWave-NOMA.

Again assume that only the users that fall into the sector \mathcal{D}_θ will participate in the NOMA transmission. Assume that there are K users in this sector. Since the users' distances are known, the base station will order the users according to the following criterion:

$$d_1 \leq \dots \leq d_K, \quad (32)$$

instead of using the effective channel gains which are not known to the base station. Similarly to the previous section, we schedule user i and user j for the NOMA transmission to act as the weak and strong users, respectively. Since a user with a shorter distance has a stronger channel condition, we take $i > j$.

Note that the density functions of the ordered distances have been found in [30] when the users are distributed randomly in a ball. The shape of the addressed area is a sector, but the steps provided in [30] are still applicable, as shown in the following. Particularly, the CDF of d_k can be calculated from the probability of the event that there are less than k users inside a sector with radius r , i.e.,

$$\begin{aligned} F_{d_k}(r) &= 1 - \sum_{i=0}^{k-1} \text{P}(E_i) \\ &= 1 - \sum_{i=0}^{k-1} e^{-\mu_{\Phi_2}(\mathcal{A}(r))} \frac{(\mu_{\Phi_2}(\mathcal{A}(r)))^i}{i!}, \end{aligned} \quad (33)$$

where $\mathcal{A}(r)$ denotes a sector with radius of r , and E_i denotes the event that there are i users in $\mathcal{A}(r)$. Following steps similar to those for obtaining (20), the factor $\mu_{\Phi_2}(\mathcal{A}(r))$ can be found as follows:

$$\mu_{\Phi_2}(\mathcal{A}(r)) = 2\Delta\lambda\phi^{-2}\gamma(2, r\phi). \quad (34)$$

Substituting the expression for $\mu_{\Phi_2}(\mathcal{A}(r))$ into the CDF expression, the CDF of d_k can be expressed as follows:

$$F_{d_k}(r) = 1 - \sum_{i=0}^{k-1} e^{-2\Delta\lambda\phi^{-2}\gamma(2, r\phi)} \frac{(2\Delta\lambda\phi^{-2}\gamma(2, r\phi))^i}{i!}. \quad (35)$$

As a result, the corresponding pdf for the k -th smallest distance can be found as follows:

$$f_{d_k}(r) = 2\Delta\lambda e^{-r\phi} r e^{-2\Delta\lambda\phi^{-2}\gamma(2, r\phi)} \frac{(2\Delta\lambda\phi^{-2}\gamma(2, r\phi))^{k-1}}{(k-1)!}, \quad (36)$$

where we have used the fact that

$$\frac{d\gamma(2, r\phi)}{dr} = e^{-r\phi} r \phi^2.$$

The difference between the above pdf expression and the one in [30] is due to the facts that the area for the addressed problem is not a ball and the addressed density is a function of r .

On the other hand, note that the angle of user k 's channel vector is independent of its distance, and it is uniformly distributed between $(\theta - \Delta)$ and $(\theta + \Delta)$. Therefore the CDF of user k 's channel gain can be obtained as follows:

$$\begin{aligned} F_k(y) &= \int_{\mathcal{D}_\theta} \left(1 - e^{-\frac{y(1+r(x)^\alpha)}{F_M(\pi|\bar{\theta}-\theta_{\pi(j)}|)}} \right) p_{X_{\pi(j)}}(x) dx \\ &= \int_{\bar{\theta}-\Delta}^{\bar{\theta}+\Delta} \int_0^{R_{\mathcal{D}}} \left(1 - e^{-\frac{y(1+r^\alpha)}{F_M(\pi|\bar{\theta}-\theta|)}} \right) \frac{f_{d_k}(r)}{2\Delta} dr d\theta. \end{aligned} \quad (37)$$

It is important to point out that the above CDF is valid only if we can find the k -th nearest node. Or in other words, if there is no boundary to \mathcal{D}_θ and the nodes are spread throughout of the plane, the above CDF can be applied. For the addressed scenario, the users are confined in \mathcal{D}_θ , i.e., $r \leq R_{\mathcal{D}}$, which means that it is possible that the k -th nearest node does not exist, i.e., there are fewer than $(k-1)$ nodes in \mathcal{D}_θ . By using the result in (37) and also considering the possible choices for the number of users in \mathcal{D}_θ , we can obtain the following lemma for the outage probability and the sum rate.

Lemma 3. *When only the users' distance information is available, the outage probability for the k -th nearest node can be written as follows:*

$$F_k^o = \sum_{n=0}^{k-1} \text{P}(K = n) + \left(1 - \sum_{n=0}^{k-1} \text{P}(K = n) \right) F_k(\eta_k), \quad (38)$$

where $k \in \{i, j\}$. Moreover, the outage sum rate can be shown as follows:

$$R_{sum}^{NOMA} = (1 - F_j(\eta_j))R_j + (1 - F_i(\eta_i))R_i, \quad (39)$$

where the k -th nearest user has a targeted data rate of R_k .

Remark 1: It is important to point out that the sum rate in (39) means that no transmission will take place if the i -th nearest user cannot be found in \mathcal{D}_θ , and the NOMA transmission is adopted even if the j -th nearest user can be found but the i -th one cannot. Note that other transmission strategies can also be used for these trivial cases which happen with low probabilities in a densely deployed network.

Remark 2: Note that one can also use a CDF expression conditioned on K to find the outage probability, but this is difficult to evaluate since the conditioning on K converts the Poisson point process to a Bernoulli one to which the result in (36) is not applicable.

Asymptotic performance analysis: While the expressions for the outage probability and the sum rate in Lemma 3 can be calculated numerically, approximations are still desirable in order to obtain greater insight. Following steps similar to those in the previous section, i.e., when Δ approaches zero, the Fejér kernel can be simplified, which yields the following approximation:

$$F_k(y) \approx \int_{\bar{\theta}-\Delta}^{\bar{\theta}+\Delta} \int_0^{R_{\mathcal{D}}} \left(1 - e^{-\frac{y(1+r^\alpha)}{M \left(1 - \frac{\pi^2 M^2 (\bar{\theta}-\theta)^2}{12} \right)}} \right) \times f_{d_k}(r) dr \frac{1}{2\Delta} d\theta. \quad (40)$$

Furthermore notice that both η_i and η_j approach zero at high SNR, which yields the following approximation:

$$F_i(\eta_i) \approx \int_{\bar{\theta}-\Delta}^{\bar{\theta}+\Delta} \int_0^{R_{\mathcal{D}}} \left(\frac{\eta_i(1+r^\alpha)}{M \left(1 - \frac{\pi^2 M^2 (\bar{\theta}-\theta)^2}{12} \right)} \right) \frac{f_{d_i}(r)}{2\Delta} dr d\theta \\ \approx \int_0^{R_{\mathcal{D}}} \left(\frac{\eta_i(1+r^\alpha)}{M} \right) f_{d_i}(r) dr \frac{1}{2\Delta} \\ \times \int_{\bar{\theta}-\Delta}^{\bar{\theta}+\Delta} \left(1 + \frac{\pi^2 M^2 (\bar{\theta}-\theta)^2}{12} \right) d\theta, \quad (41)$$

and $F_j(\eta_j)$ can be obtained similarly. The integral over θ can be obtained by following steps similar to those in the previous section. In addition, define the integral over r , a factor not related to the transmit SNR, as follows:

$$Q_{1,i} \triangleq \int_0^{R_{\mathcal{D}}} \left(\frac{(1+r^\alpha)}{M} \right) \lambda e^{-r\phi} r e^{-2\Delta\lambda\phi^{-2}\gamma(2,r\phi)} \times \frac{(2\Delta\lambda\phi^{-2}\gamma(2,r\phi))^{i-1}}{(i-1)!} dr. \quad (42)$$

Therefore the outage probability can be approximated as follows:

$$F_i(\eta_i) \approx \left(1 + \frac{\pi^2 M^2 \Delta^2}{36} \right) Q_{1,i} \eta_i,$$

which demonstrates that the use of distance information only yields a diversity gain of one for all the users. This is expected since the base station has access to partial CSI only and the dynamics of the fading gains cannot be used.

B. With One-Bit Feedback

In the case in which the number of users in the sector \mathcal{D}_θ is very large, feeding these users' effective channel gains or distances back to the base station can still be very demanding. As an alternative, asking each user to feed only one bit about its channel quality back to the base station can substantially reduce the system overhead.

In particular, the base station will first set a threshold, ξ , $\xi > 0$, which will be broadcast to all the users prior to the downlink transmission. Each user in the sector will compare its effective channel gain with ξ and send 1 to the base station if its channel gain is larger than ξ , otherwise it will send 0 to the base station. As a result, the users in the sector will be divided into two sets, denoted by \mathcal{S}_1 and \mathcal{S}_2 , respectively.

Particularly, the users in \mathcal{S}_2 are the ones which feed 1 back to the base station, i.e.,

$$\mathcal{S}_2 \triangleq \{i | x_i \in \mathcal{D}_\theta, |\mathbf{h}_i^H \mathbf{p}|^2 > \xi\}, \quad (43)$$

and \mathcal{S}_1 is defined similarly by grouping those users whose feedbacks are 0.

When there is more than one user in \mathcal{D}_θ , i.e., $K \geq 2$, and $|\mathcal{S}_n| \neq 0$, $n \in \{1, 2\}$, the base station will randomly select one user from \mathcal{S}_1 to be paired with another user randomly selected from \mathcal{S}_2 . If all the K nodes are in one group, the base station will randomly select two users from this group for the implementation of NOMA. If there is only one user in the sector, i.e., $K = 1$, this user will be served solely by the base station. No user will be served if both sets are empty, which happens only if $K = 0$. In the following, we will focus on the case with $K \geq 2$.

The following lemma provides the outage probabilities for the users selected to act as the NOMA strong and weak users, respectively.

Lemma 4. Suppose that there are $K \geq 2$ users in \mathcal{D}_θ . When each user only feeds one bit back to the base station using the above protocol, the outage probability for the user selected to act as the weak user is given by

$$P_{\mathcal{S}_1}^o = F_{\mathcal{S}_1|K}(\tilde{\eta}_1) \sum_{n=1}^K \text{P}(|\mathcal{S}_1| = n) + \text{P}(|\mathcal{S}_1| = 0) F_{\mathcal{S}_2|K}(\tilde{\eta}_1), \quad (44)$$

and the outage probability for the user selected to act as the strong user is given by

$$P_{\mathcal{S}_2}^o = F_{\mathcal{S}_2|K}(\tilde{\eta}_2) \sum_{n=1}^K \text{P}(|\mathcal{S}_2| = n) + \text{P}(|\mathcal{S}_2| = 0) F_{\mathcal{S}_1|K}(\tilde{\eta}_2), \quad (45)$$

where $\text{P}(|\mathcal{S}_2| = n) = \binom{K}{n} (F_{\pi(j)}(\xi))^{K-n} (1 - F_{\pi(j)}(\xi))^n$, $\text{P}(|\mathcal{S}_1| = n) = \text{P}(|\mathcal{S}_2| = K - n)$, $F_{\mathcal{S}_1|K}(y) = \frac{F_{\pi(j)}(\min\{y, \xi\})}{F_{\pi(j)}(\xi)}$, $F_{\mathcal{S}_2|K}(y) = \max\left\{0, \frac{F_{\pi(j)}(y) - F_{\pi(j)}(\xi)}{1 - F_{\pi(j)}(\xi)}\right\}$, $\tilde{\eta}_2 = \max\left\{\frac{\tilde{\epsilon}_1}{\beta_1^2 - \beta_2^2 \tilde{\epsilon}_1}, \frac{\tilde{\epsilon}_2}{\rho \beta_2^2}\right\}$, $\tilde{\eta}_1 = \frac{\tilde{\epsilon}_1}{\beta_1^2 - \beta_2^2 \tilde{\epsilon}_1}$, $\tilde{\epsilon}_k = 2^{\tilde{R}_k} - 1$, for $k \in \{1, 2\}$, and \tilde{R}_1 and \tilde{R}_2 denote the targeted rates for the users selected to act as the weak and strong users, respectively.

Proof: Since there are $K \geq 2$ users in the sector, the outage probability experienced by the user chosen to act as the strong user in NOMA can be expressed as follows:

$$P_{\mathcal{S}_2}^o = F_{\mathcal{S}_2|K}(\tilde{\eta}_2) \text{P}(|\mathcal{S}_2| \text{ is not empty}) + \text{P}(|\mathcal{S}_2| \text{ is empty}) F_{\mathcal{S}_1|K}(\tilde{\eta}_2), \quad (46)$$

where $F_{\mathcal{S}_2|K}(\cdot)$ denotes the CDF of the effective channel gain of a user randomly selected from \mathcal{S}_2 and its expression will be evaluated later. The probability $\text{P}(|\mathcal{S}_2| \text{ is not empty})$ is equivalent to $\sum_{n=1}^K \text{P}(|\mathcal{S}_1| = n)$. Note that $F_{\mathcal{S}_1|K}(\tilde{\eta}_2)$, the CDF of the weak user's channel, is used for the case of $|\mathcal{S}_2| = 0$ since the base station will select one user randomly from \mathcal{S}_1 to act as the strong user with the targeted rate of \tilde{R}_2 for the NOMA transmission. Similarly, the outage probability

experienced by the user selected to act as the weak user can be expressed as follows:

$$P_{\mathcal{S}_1}^o = F_{\mathcal{S}_1|K}(\tilde{\eta}_1) \sum_{n=1}^K P(|\mathcal{S}_1| \text{ is not empty}) \quad (47)$$

$$+ P(|\mathcal{S}_1| \text{ is empty}) F_{\mathcal{S}_2|K}(\tilde{\eta}_1),$$

where the variables are defined similarly to their counterparts in (46).

Given that there are K users in the sector, the probability for the case of $|\mathcal{S}_2| = n$ can be obtained as shown in the lemma, which is due to the fact that all the users' channels are independent and identically distributed.

The CDF of the effective channel gain of a user randomly selected from \mathcal{S}_1 can be expressed as follows:

$$F_{\mathcal{S}_1|K}(y) = P(|\mathbf{h}_j^H \mathbf{p}|^2 < y \mid X_j \in \mathcal{D}_\theta, j \in \mathcal{S}_1) \quad (48)$$

$$= \frac{\int_{\mathcal{D}_\theta} P(|\mathbf{h}_j^H \mathbf{p}|^2 < \min\{y, \xi\} \mid X_j = x_j) dP_{X_j}}{\int_{\mathcal{D}_\theta} P(|\mathbf{h}_j^H \mathbf{p}|^2 < \xi \mid X_j = x_j) dP_{X_j}}.$$

Following steps similar to those in Section III, the addressed CDF can be obtained as follows:

$$F_{\mathcal{S}_1|K}(y) = \frac{F_{\pi(j)}(\min\{y, \xi\})}{F_{\pi(j)}(\xi)}, \quad (49)$$

for $\xi > 0$.

On the other hand, the CDF of the effective channel gain of a user randomly selected from \mathcal{S}_2 can be expressed as follows:

$$F_{\mathcal{S}_2|K}(y) = P(|\mathbf{h}_j^H \mathbf{p}|^2 < y \mid X_j \in \mathcal{D}_\theta, j \in \mathcal{S}_2) \quad (50)$$

$$= \frac{\int_{\mathcal{D}_\theta} P(\xi < |\mathbf{h}_j^H \mathbf{p}|^2 < y \mid X_j = x_j) dP_{X_j}}{\int_{\mathcal{D}_\theta} P(|\mathbf{h}_j^H \mathbf{p}|^2 > \xi \mid X_j = x_j) dP_{X_j}},$$

if $y > \xi$, otherwise $F_{\mathcal{S}_2|K}(y) = 0$. Again following steps similar to those in Section III, this CDF can be found as follows:

$$F_{\mathcal{S}_2|K}(y) = \frac{F_{\pi(j)}(y) - F_{\pi(j)}(\xi)}{1 - F_{\pi(j)}(\xi)}, \quad (51)$$

if $\xi < y$, otherwise $F_{\mathcal{S}_2|K}(y) = 0$. Substituting (49) and (51) into (47) and (46), the outage probabilities in the lemma can be obtained and the proof is complete. ■

By using the outage probabilities obtained in the above lemma, one can easily find an expression for the outage sum rate, which is omitted here due to space limitations.

Obviously the choice of ξ will have a significant impact on the performance of the addressed one-bit feedback scenario. To investigate this impact, we will first study the impact of ξ on the CDFs, $F_{\mathcal{S}_1|K}(\tilde{\eta}_1)$ and $F_{\mathcal{S}_2|K}(\tilde{\eta}_2)$.

1) *The impact of the threshold on $F_{\mathcal{S}_k|K}(\tilde{\eta}_k)$* : Because $\tilde{\eta}_1$ approaches zero at high SNR, $\min\{\tilde{\eta}_1, \xi\}$ will also approach zero at high SNR, with a rate of decaying no smaller than $\tilde{\eta}_1$. Note that the outage probability of the user selected to act as the NOMA weak user is related to $F_{\mathcal{S}_1|K}(\tilde{\eta}_1)$, which can be

approximated at high SNR as follows:

$$F_{\mathcal{S}_1|K}(\tilde{\eta}_1) = \frac{F_{\pi(j)}(\min\{\tilde{\eta}_1, \xi\})}{F_{\pi(j)}(\xi)} \quad (52)$$

$$\approx \frac{\min\{\tilde{\eta}_1, \xi\}}{2M\gamma(2, R_{\mathcal{D}}\phi)F_{\pi(j)}(\xi)} \left(2 + \frac{\pi^2 M^2 \Delta^2}{18}\right)$$

$$\times (\gamma(2, R_{\mathcal{D}}\phi) + \phi^{-\alpha}\gamma(\alpha + 2, R_{\mathcal{D}}\phi)).$$

In this paper, we are interested in the following two choices of ξ .

- If ξ is a constant and not a function of the transmit SNR, ρ , the following holds at high SNR:

$$F_{\mathcal{S}_1|K}(\tilde{\eta}_1) \sim \frac{1}{\rho}. \quad (53)$$

- If ξ decreases at a rate of $\frac{1}{\rho^x}$, i.e., $\xi \sim \frac{1}{\rho^x}$, $x > 0$, we have the following approximation:

$$F_{\mathcal{S}_1|K}(\tilde{\eta}_1) \approx \frac{\min\{\tilde{\eta}_1, \xi\}}{\xi}. \quad (54)$$

On the other hand, the impact of ξ on $F_{\mathcal{S}_2|K}(\tilde{\eta}_2)$ can be demonstrated as follows.

- If ξ is a constant, $F_{\mathcal{S}_2|K}(\tilde{\eta}_2) = 0$ at high SNR, since $\tilde{\eta}_2$ approaches zero at high SNR and hence

$$P(\xi < |\mathbf{h}_j^H \mathbf{p}|^2 < \tilde{\eta}_2 \mid X_j = x_j) = 0.$$

- If $\xi \sim \frac{1}{\rho^x}$, $x > 0$, we have the following approximation:

$$F_{\mathcal{S}_2|K}(\tilde{\eta}_2) = \frac{F_{\pi(j)}(\tilde{\eta}_2) - F_{\pi(j)}(\xi)}{1 - F_{\pi(j)}(\xi)} \quad (55)$$

$$\approx \frac{(\tilde{\eta}_2 - \xi)}{2M\gamma(2, R_{\mathcal{D}}\phi)} \left(2 + \frac{\pi^2 M^2 \Delta^2}{18}\right)$$

$$\times (\gamma(2, R_{\mathcal{D}}\phi) + \phi^{-\alpha}\gamma(\alpha + 2, R_{\mathcal{D}}\phi)),$$

if $\xi < \tilde{\eta}_2$, otherwise $F_{\mathcal{S}_2|K}(\tilde{\eta}_2) = 0$.

2) *The impact of ξ on the users' outage probabilities, $P_{\mathcal{S}_k}^o$* : We first focus on the user selected to act as the weak user, whose diversity gain is shown in the following lemma.

Lemma 5. *For the two considered choices of the threshold, i.e., either $\xi \sim \frac{1}{\rho^x}$ or ξ is a constant, the diversity order of the user selected to act as the weak user is always one.*

Proof: For notational simplicity, let $c_2 = \frac{(2 + \frac{\pi^2 M^2 \Delta^2}{18})}{2M\gamma(2, R_{\mathcal{D}}\phi)} (\gamma(2, R_{\mathcal{D}}\phi) + \phi^{-\alpha}\gamma(\alpha + 2, R_{\mathcal{D}}\phi))$. It can be shown that the probability of having n users in group \mathcal{S}_2 can be approximated as follows:

$$P(|\mathcal{S}_2| = n) = \binom{K}{n} (F_{\pi(j)}(\xi))^{K-n} (1 - F_{\pi(j)}(\xi))^n \quad (56)$$

$$\approx \binom{K}{n} c_2^{K-n} \xi^{K-n},$$

if ξ approaches zero. If ξ is not a function of ρ , neither is this probability.

- If ξ is a constant, the outage probability of the user selected to act as the weak user can be simplified as follows:

$$P_{\mathcal{S}_1}^o = F_{\mathcal{S}_1|K}(\tilde{\eta}_1) \sum_{n=1}^K P(|\mathcal{S}_1| = n) \sim \frac{1}{\rho}, \quad (57)$$

since $F_{\mathcal{S}_2|K}(\tilde{\eta}_1) = 0$, $\sum_{n=1}^K P(|\mathcal{S}_1| = n)$ is a constant and $F_{\mathcal{S}_1|K}(\tilde{\eta}_1) \sim \frac{1}{\rho}$ as explained in (53). Therefore the user's diversity order is one for this choice of ξ .

- If $\xi \sim \frac{1}{\rho^x}$, $x > 0$, the outage probability for the user selected to act as the weak user can be approximated as follows:

$$\begin{aligned} P_{\mathcal{S}_1}^o &= F_{\mathcal{S}_1|K}(\tilde{\eta}_1) \sum_{n=1}^K P(|\mathcal{S}_1| = n) \\ &\quad + P(|\mathcal{S}_1| = 0) F_{\mathcal{S}_2|K}(\tilde{\eta}_1) \\ &\approx \frac{\min\{\tilde{\eta}_1, \xi\}}{\xi} \sum_{n=1}^K \binom{K}{n} c_2^n \xi^n + c_2 \max\{0, \tilde{\eta}_1 - \xi\} \\ &\approx \frac{\min\{\tilde{\eta}_1, \xi\}}{\xi} K c_2 \xi + c_2 \max\{0, \tilde{\eta}_1 - \xi\}, \end{aligned} \quad (58)$$

which is always at the order of $\frac{1}{\rho}$ as explained in the following.

If $x > 1$, $\min\{\tilde{\eta}_1, \xi\} = \xi$ and $\max\{0, \tilde{\eta}_1 - \xi\} \approx \tilde{\eta}_1$. These two observations lead to the following approximation:

$$P_{\mathcal{S}_1}^o \approx \xi K c_2 + c_2 \tilde{\eta}_1 \sim \frac{1}{\rho}, \quad (59)$$

since $\tilde{\eta}_1$ is dominant.

If $x = 1$, we have the following approximation:

$$P_{\mathcal{S}_1}^o \approx \min\{\tilde{\eta}_1, \xi\} K c_2 + c_2 \max\{0, \tilde{\eta}_1 - \xi\} \sim \frac{1}{\rho}, \quad (60)$$

since both $\min\{\tilde{\eta}_1, \xi\}$ and $|\tilde{\eta}_1 - \xi|$ are at the order of $\frac{1}{\rho}$. Further, if $0 < x < 1$, we have the following approximation:

$$P_{\mathcal{S}_1}^o \approx \tilde{\eta}_1 K c_2 \sim \frac{1}{\rho}. \quad (61)$$

Therefore, we can conclude that, as long as $\xi \sim \frac{1}{\rho^x}$, $x > 0$, the diversity order of the user selected to act as the weak user is one.

Since the user's diversity order is one for both cases, the proof is complete. ■

However, the impact of ξ on $P_{\mathcal{S}_2}^o$ is more complicated as illustrated in the following:

- If ξ is a constant, the diversity order of the user selected to act as the strong user is one, since

$$P_{\mathcal{S}_2}^o = P(|\mathcal{S}_2| = 0) F_{\mathcal{S}_1|K}(\tilde{\eta}_2) \sim \frac{1}{\rho},$$

which is due to the following facts: $F_{\mathcal{S}_2|K}(\tilde{\eta}_2) = 0$, $P(|\mathcal{S}_2| = 0)$ is a constant and $F_{\mathcal{S}_1|K}(\tilde{\eta}_2) \sim \frac{1}{\rho}$ as explained in (53).

- If $\xi \sim \frac{1}{\rho^x}$, $x \geq 1$, the outage probability for the user selected to act as the strong user can be approximated as follows:

$$\begin{aligned} P_{\mathcal{S}_2}^o &= F_{\mathcal{S}_2|K}(\tilde{\eta}_2) \sum_{n=1}^K P(|\mathcal{S}_2| = n) \\ &\quad + P(|\mathcal{S}_2| = 0) F_{\mathcal{S}_1|K}(\tilde{\eta}_2) \\ &\approx c_2 \max\{0, \tilde{\eta}_2 - \xi\} + c_2^K \xi^K \frac{\min\{\tilde{\eta}_2, \xi\}}{\xi}. \end{aligned} \quad (62)$$

As can be seen from (62), the choice of the threshold ξ has significant impact on the achievable diversity gain. For example, a full diversity gain of K can be obtained by using the following choice of ξ :

$$\xi = \tilde{\eta}_2 - \frac{1}{\rho^K}. \quad (63)$$

V. RANDOM BEAMFORMING: A MULTIPLE-BEAM CASE

A. System Model and Outage Performance

Consider a scenario in which the base station will form N , $1 < N \leq M$, orthonormal beams, denoted by \mathbf{p}_m , $1 \leq m \leq N$, where $\mathbf{p}_m^H \mathbf{p}_m = 1$ and $\mathbf{p}_m^H \mathbf{p}_n = 0$ if $m \neq n$. These beamforming vectors are predefined, and it is assumed that they are known to the base station and the users prior to transmission. Following [13] and [23], these N orthonormal beamforming vectors can be constructed as follows:

$$\mathbf{p}_m = \mathbf{a} \left(\zeta + \frac{2(m-1)}{N} \right), \quad (64)$$

for $1 \leq m \leq N$, where ζ denotes a random variable following a uniform distribution between -1 and 1 . For notational simplicity, we denote $\zeta + \frac{2(m-1)}{N}$ by θ_m . Again this beamformer can also be viewed as a special case of the hybrid precoding design in [24], in which the fully-connected architecture is used with N radio frequency chains, M antennas and a digital precoding matrix set as an identity matrix.

Prior to downlink transmission, the base station will first broadcast pilot signals on these N orthogonal beams. Similarly to \mathcal{D}_θ , define \mathcal{D}_{θ_m} as the wedge-shaped sector around θ_m with a central angle of 2Δ , as shown in Fig. 1. Only the users that fall into the sector \mathcal{D}_{θ_m} will participate in the NOMA transmission on beam m . Denote the number of users in \mathcal{D}_{θ_m} by K_m and the k^{th} user's channel by $\mathbf{h}_{m,k}$, $1 \leq k \leq K_m$. Each user will measure its effective channel gain on its corresponding beam, where user k 's effective channel gain on the m -th beam is given by $|\mathbf{h}_{m,k}^H \mathbf{p}_m|^2$. Without loss of generality, we assume that the base station schedules user i and user j on beam m , to act as the weak and strong users, respectively.

Therefore, the base station will superimpose two users' messages on each of the N beams as follows:

$$\sum_{m=1}^N \mathbf{p}_m (\beta_{m,1} s_{m,i} + \beta_{m,2} s_{m,j}), \quad (65)$$

where $\beta_{m,1}^2 + \beta_{m,2}^2 = 1$.

Therefore, user j on beam m will receive the following observation:

$$\begin{aligned} y_{m,j} &= \mathbf{h}_{m,j}^H \sum_{n=1}^N \mathbf{p}_n (\beta_{n,1} s_{n,i} + \beta_{n,2} s_{n,j}) + n_{m,j} \\ &= \mathbf{h}_{m,j}^H \mathbf{p}_m (\beta_{m,1} s_{m,i} + \beta_{m,2} s_{m,j}) \\ &\quad + \mathbf{h}_{m,j}^H \sum_{n=1, n \neq m}^N \mathbf{p}_n (\beta_{n,1} s_{n,i} + \beta_{n,2} s_{n,j}) + n_{m,j}, \end{aligned} \quad (66)$$

where $n_{m,j}$ denotes additive Gaussian noise.

User j on beam m will first decode the message to user i in the same pair, and then remove this message from its observation. Such SIC needs to be carried out before its own message is decoded. As a result, the SINR for user j on beam m to decode its partner's message can be expressed as follows:

$$\text{SINR}_{m,i \rightarrow j} = \frac{|\mathbf{h}_{m,j}^H \mathbf{p}_m|^2 \beta_{m,1}^2}{|\mathbf{h}_{m,j}^H \mathbf{p}_m|^2 \beta_{m,2}^2 + \sum_{n \neq m} |\mathbf{h}_{m,j}^H \mathbf{p}_n|^2 + \frac{1}{\rho}}. \quad (67)$$

Define $R_{m,1}$ as the targeted rate for user i on beam m and $\epsilon_{m,1} = 2^{R_{m,1}} - 1$, where $\epsilon_{m,2}$ and $R_{m,2}$ are defined for user j similarly. If $\text{SINR}_{m,i \rightarrow j} \geq \epsilon_{m,1}$, intra-group interference can be cancelled and the user can decode its own information with the following SINR:

$$\text{SINR}_{m,j} = \frac{|\mathbf{h}_{m,j}^H \mathbf{p}_m|^2 \beta_{m,2}^2}{\sum_{n \neq m} |\mathbf{h}_{m,j}^H \mathbf{p}_n|^2 + \frac{1}{\rho}}. \quad (68)$$

User i on beam m will decode its own message directly with the following SINR:

$$\text{SINR}_{m,i} = \frac{|\mathbf{h}_{m,i}^H \mathbf{p}_m|^2 \beta_{m,1}^2}{|\mathbf{h}_{m,i}^H \mathbf{p}_m|^2 \beta_{m,2}^2 + \sum_{n \neq m} |\mathbf{h}_{m,i}^H \mathbf{p}_n|^2 + \frac{1}{\rho}}. \quad (69)$$

Different from the case with one beam, the users' SINRs are functions not only of $|\mathbf{h}_{m,i}^H \mathbf{p}_m|^2$ but also of $|\mathbf{h}_{m,i}^H \mathbf{p}_n|^2$, $n \neq m$. In conventional non-NOMA scenarios, users can be scheduled according to their SINRs, i.e., the user with the strongest SINR on beam m will be selected to be served on this beam. However, in the addressed scenario, one user can have two different SINR functions. For example, user j 's performance depends on two different SINR functions, $\text{SINR}_{m,i \rightarrow j}$ and $\text{SINR}_{m,j}$. For the purpose of illustration, we focus on a simple user scheduling scheme based on distances, a strategy similar to the one proposed in Section IV-A. Therefore we can order these users who will participate in the NOMA transmission on beam m as follows:

$$d_{m,1} \leq \dots \leq d_{m,K_m}. \quad (70)$$

Furthermore suppose that user i has a distance larger than that of user j , i.e., $i > j$.

The outage probability experienced by user j can be expressed as follows:

$$P_{m,j}^o = 1 - \text{P}(\text{SINR}_{m,i \rightarrow j} > \epsilon_{m,1}, \text{SINR}_{m,j} > \epsilon_{m,2}). \quad (71)$$

Again applying the mmWave channel model, $\text{SINR}_{m,i \rightarrow j}$, can be written as follows:

$$\begin{aligned} \text{SINR}_{m,i \rightarrow j} &= \frac{|a_{m,j}|^2}{(1+d_{m,j}^\alpha)} F_M(\pi[\bar{\theta}_m - \theta_{m,j}]) \beta_{m,1}^2 \\ &\times \left(\frac{|a_{m,j}|^2}{(1+d_{m,j}^\alpha)} F_M(\pi[\bar{\theta}_m - \theta_{m,j}]) \beta_{m,2}^2 \right. \\ &\left. + \sum_{n \neq m} \frac{|a_{m,j}|^2}{(1+d_{m,j}^\alpha)} F_M(\pi[\bar{\theta}_n - \theta_{m,j}]) + \frac{1}{\rho} \right)^{-1}. \end{aligned} \quad (72)$$

Similarly, $\text{SINR}_{m,j}$, can be expressed as follows:

$$\text{SINR}_{m,j} = \frac{\frac{|a_{m,j}|^2}{(1+d_{m,j}^\alpha)} F_M(\pi[\bar{\theta}_m - \theta_{m,j}]) \beta_{m,2}^2}{\sum_{n \neq m} \frac{|a_{m,j}|^2}{(1+d_{m,j}^\alpha)} F_M(\pi[\bar{\theta}_n - \theta_{m,j}]) + \frac{1}{\rho}}. \quad (73)$$

Unlike those SINR functions in the previous sections, the SINRs for the case with multiple beams become more complicated. An interesting observation is that the three factors in the numerator and denominator of $\text{SINR}_{m,i \rightarrow j}$ share the same fading coefficient. In this case, the outage probability of user j on beam m can be expressed as shown in (74) at the top of the following page, if $F_{j,m}^m \beta_{m,1}^2 > \epsilon_{m,1} F_{j,m}^m \beta_{m,2}^2 + \sum_{n \neq m} \epsilon_{m,1} F_{j,n}^m$ and $F_{j,m}^m \beta_{m,2}^2 > \sum_{n \neq m} \epsilon_{m,2} F_{j,n}^m$, otherwise the outage probability will be always one, where $F_{j,n}^m \triangleq F_M(\pi[\bar{\theta}_n - \theta_{m,j}])$. The outage probability at user i can be obtained similarly.

B. Asymptotic Performance Analysis

Without loss of generality, we focus on the first beam, i.e., $m = 1$. In this case, the factor $F_{j,n}^m$ can be written as follows:

$$\begin{aligned} F_{j,n}^1 &= F_M(\pi[\theta_{1,j} - \bar{\theta}_n]) \\ &= F_M\left(\pi\left[\theta_{1,j} - \bar{\theta}_1 - \frac{2(n-1)}{N}\right]\right), \end{aligned} \quad (75)$$

where $2 \leq n \leq N$. We have the following Taylor series approximation:

$$F_{j,n}^1 = \sum_{l=0}^{\infty} F_M^{(l)}\left(-\frac{2(n-1)\pi}{N}\right) \frac{(\theta_{1,j} - \bar{\theta}_1)^l}{l!}, \quad (76)$$

where $F_M^{(l)}(x)$ denotes the l -th derivative of $F_M(x)$. Here we assume that the derivatives, $F_M^{(l)}\left(-\frac{2(n-1)\pi}{N}\right)$, exist for all orders. Assume that the beams are separated with sufficient gaps, and one can expect that $F_M\left(-\frac{2(n-1)\pi}{N}\right) \rightarrow 0$, for $2 \leq n \leq N$. Further assuming $\Delta \rightarrow 0$, $(\bar{\theta}_n - \theta_{1,j})$ approaches zero, which means

$$\begin{aligned} F_{j,n}^1 &\approx F_M\left(-\frac{2(n-1)\pi}{N}\right) \\ &+ F_M^{(1)}\left(-\frac{2(n-1)\pi}{N}\right) (\bar{\theta}_1 - \theta_{1,j}), \end{aligned} \quad (77)$$

where $F_M^{(1)}(x) = \frac{\sin Mx}{1-\cos(x)} - \frac{(1-\cos Mx)\sin x}{M(1-\cos x)^2}$.

Therefore the sum of the interference terms in the SINR expressions can be approximated as follows:

$$\sum_{n \neq 1} F_{j,n}^1 \approx c_2 + c_3(\bar{\theta}_1 - \theta_{1,j}), \quad (78)$$

where $c_2 = \sum_{n \neq 1} F_M\left(-\frac{2(n-1)\pi}{N}\right)$ and $c_3 = \sum_{n \neq 1} F_M^{(1)}\left(-\frac{2(n-1)\pi}{N}\right)$. For the case $n = 1$, we have $F_{j,1}^1 \approx M$, (79)

$$P_{m,j}^o = 1 - \int_{\bar{\theta}-\Delta}^{\bar{\theta}+\Delta} \int_0^{R_D} \exp\left\{-\max\left\{\frac{\epsilon_{m,1}(1+d_{m,j}^\alpha)}{F_{j,m}^m \beta_{m,1}^2 - \epsilon_{m,1} F_{j,m}^m \beta_{m,2}^2 - \sum_{n \neq m} \epsilon_{m,1} F_{j,n}^m}, \frac{\epsilon_{m,2}(1+d_{m,j}^\alpha)}{F_{j,m}^m \beta_{m,2}^2 - \sum_{n \neq m} \epsilon_{m,2} F_{j,n}^m}\right\}\right\} f_{d_j}(r) dr d\theta, \quad (74)$$

which is obtained from (27). As a result, at high SNR, the outage probability experienced by user i can be expressed as follows:

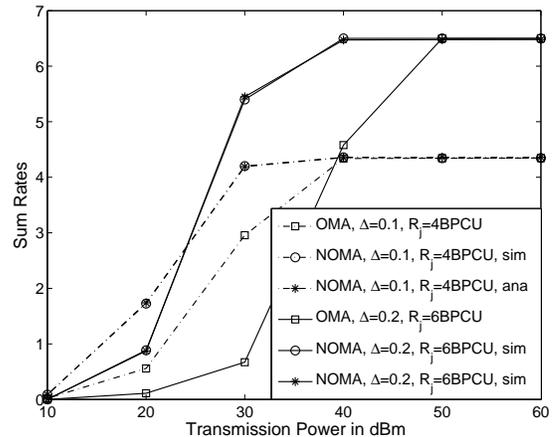
$$\begin{aligned} P_{1,i}^o &\approx \frac{1}{2\Delta} \int_{\bar{\theta}-\Delta}^{\bar{\theta}+\Delta} \int_0^{R_D} \frac{\epsilon_{1,1}(1+d_{1,i}^\alpha)}{F_{i,1}^1 \beta_{1,1}^2 - \epsilon_{1,1} F_{i,1}^1 \beta_{1,2}^2 - \sum_{n \neq 1} \epsilon_{1,1} F_{i,n}^1} \\ &\quad \times f_{d_i}(r) dr d\theta \\ &\approx MQ_{1,i} \int_{-\Delta}^{\Delta} \frac{\epsilon_{1,1}}{(Mc_4 - c_2 \epsilon_{1,1}) - c_3 \epsilon_{1,1} y} dy \\ &\approx \frac{2\Delta MQ_{1,i} \epsilon_{1,1}}{\rho (Mc_4 - c_2 \epsilon_{1,1})}, \end{aligned} \quad (80)$$

where $c_4 = \beta_{1,1}^2 - \epsilon_{1,1} \beta_{1,2}^2$. The outage probability for user j can be obtained similarly. As a result, following steps similar to those in Section IV-A, the outage sum rate and the outage probabilities can be obtained.

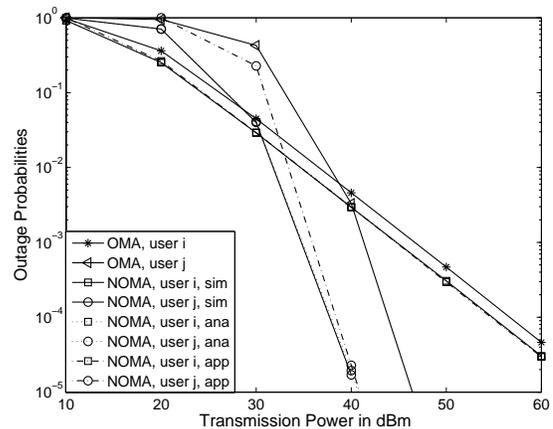
VI. NUMERICAL STUDIES

In this section, the performance of the proposed mmWave-NOMA transmission schemes are evaluated by using computer simulations, where the accuracy of the developed analytical results will also be verified. The path loss exponent is set as $\alpha = 2$, since line-of-sight links are focused. The radius of \mathcal{D} is $R_D = 10\text{m}$, the noise power is -30dBm , the blockage parameter is set as $\phi = 0.1$, and $\beta_i^2 = \frac{3}{4}$ and $\beta_j^2 = \frac{1}{4}$ are used as the NOMA power allocation coefficients. It is worth pointing out that our analytical results are developed for arbitrary choices of these parameters, and using other choices of these parameters will lead to conclusions similar to those drawn in this section.

In Fig. 2, the performance of the proposed random beamforming scheme in mmWave-NOMA systems with perfect CSI is studied, where the mmWave-OMA scheme is used as a benchmark. Fig. 2.(a) shows the outage sum rates achieved by the two MA schemes, and Fig. 2.(b) shows the outage probabilities of the two transmission schemes. As can be observed from Fig. 2.(a), the use of NOMA can yield a significant sum rate gain over the OMA scheme, and this gain increases when the targeted data rate of the strong user is increased. For example, for $R_j = 4$ bits per channel use (BPCU), the gain of mmWave-NOMA over mmWave-OMA is 1 BPCU, when the transmission power of the base station is 30 dBm. When R_j is increased to 6 BPCU, the performance gain of the NOMA scheme over OMA becomes 5 BPCU. On the other hand, Fig. 2.(b) shows that the mmWave-NOMA scheme can also effectively reduce the outage probability, compared to OMA, particularly for the user with the stronger channel. It is also important to point out that the developed approximation



(a) Sum Rates

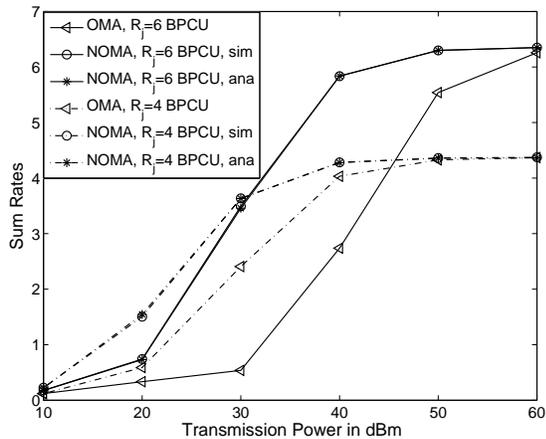


(b) Outage Probabilities $R_2 = 6$ BPCU and $K = 5$

Fig. 2. The performance of mmWave-NOMA and mmWave-OMA with perfect CSI. $M = 4$, $\lambda = 1$, $\Delta = 0.1$, $R_i = 0.5$ BPCU, $i = 1$ and $j = K$.

results for the sum rate and the outage probabilities are tight at high SNR, and the developed exact expressions match the simulation results perfectly.

In Fig. 3, the performance of the mmWave-NOMA and mmWave-OMA schemes is compared, for the situation in which the base station has access to the users' distance information only. The trivial cases in which the i -th and j -th nearest nodes do not exist can cause error floors to the outage probabilities. Therefore, we slightly change the definition of the outage probability by counting only the cases in which the two nodes can be found in \mathcal{D}_θ . Take the outage probability for user i as an example. The outage probability curves are obtained by using $\frac{n_3}{n_1 - n_2}$, where n_1 denotes the total number of simulations, n_2 denotes the number of events in which user i cannot be found in \mathcal{D}_θ , and n_3 denotes the number of outage events by excluding the outage events caused by the case in



(a) Sum Rates

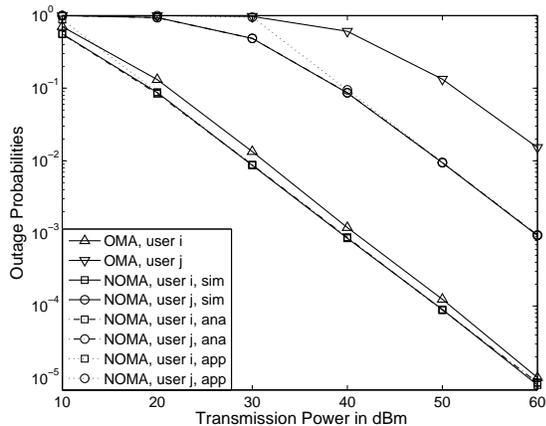
(b) Outage Probabilities $R_2 = 6$ BPCU

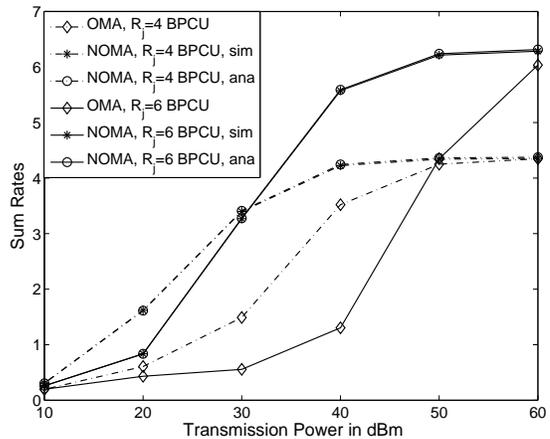
Fig. 3. The performance of mmWave-NOMA and mmWave-OMA by using the distance information only. $M = 4$, $\lambda = 1$, $\Delta = 0.1$, $R_i = 0.5$ BPCU, $i = 4$ and $j = 1$.

which user i cannot be found (i.e., n_2). This is consistent with (38) since the probability shown in the figure is equivalent to the following one

$$\frac{F_k^o - \sum_{n=0}^{k-1} P(K = n)}{1 - \sum_{n=0}^{k-1} P(K = n)}. \quad (81)$$

As can be observed from both figures, the use of NOMA can yield a significant performance gain in the sum rate and effectively reduce the outage probability, compared to the OMA scheme, even if only the distance information is available to the base station. Again both figures also demonstrate the accuracy of the developed analytical results.

Fig. 4 shows the sum rate and the outage probabilities achieved by the one-bit feedback scheme, and Fig. 5 shows the impact of the threshold ξ on the users' outage probabilities and diversity gains. Consistent with the previous figures, Fig. 4 demonstrates that the use of NOMA can significantly improve the performance of mmWave communications with random beamforming. Recall that the choice of the threshold has a significant impact on the performance of the one-bit feedback scheme. As discussed in Section IV-B, the diversity gain of the strong user is particularly sensitive to the choice of the threshold, and a choice of $\xi = \tilde{\eta}_j - \frac{1}{\rho^K}$ yields a diversity gain of K , whereas the diversity gain of the weak user is always



(a) Sum Rates

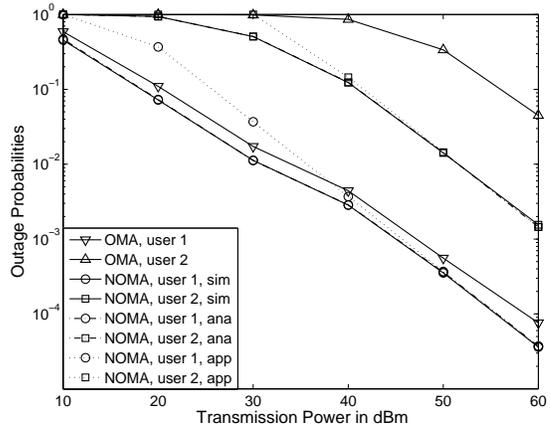
(b) Outage Probabilities $R_2 = 6$ BPCU and $K = 5$

Fig. 4. The performance of mmWave-NOMA and mmWave-OMA with one-bit feedback. $M = 4$, $\lambda = 1$, $\Delta = 0.1$, and $R_i = 0.5$ BPCU. The threshold is set as $\frac{1}{\rho}(\tilde{\eta}_1 + \tilde{\eta}_2)$.

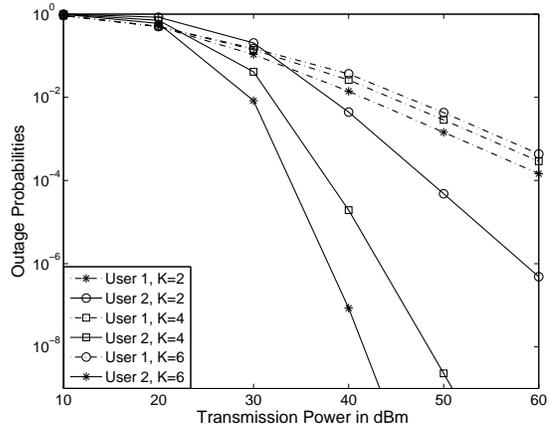
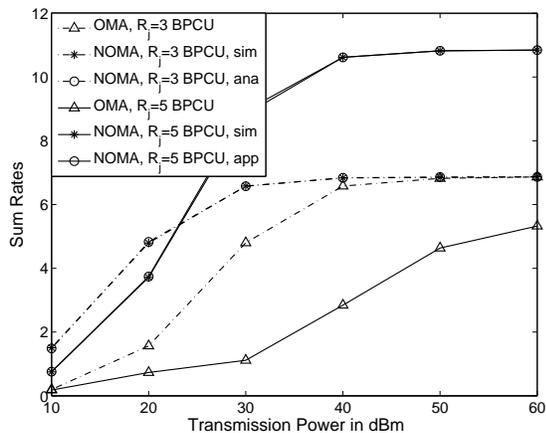


Fig. 5. The impact of the threshold on the performance of the one-bit feedback scheme. $M = 4$, $\lambda = 1$, $\Delta = 0.1$, $R_1 = 1.5$ BPCU and $R_2 = 4$ BPCU. The threshold is set as $(\tilde{\eta}_2 - \frac{1}{\rho^K})$.

one for the discussed choices of ξ . Fig. 5 clearly confirms these analytical results and demonstrates the impact of ξ on the diversity gain. For example, the slope of the strong user's outage probability curve becomes larger when increasing K , which demonstrates that the diversity gain of this user is an increasing function of K . On the other hand, the slope for the other user's outage probability curve is always the same, which shows that the diversity gain of the weak user is not



(a) Sum Rates

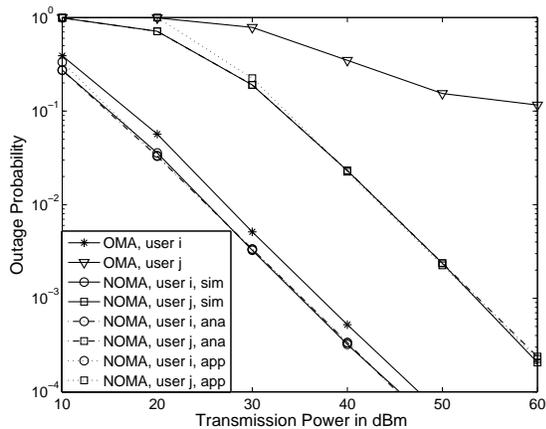
(b) Outage Probabilities $R_j = 5$ BPCU

Fig. 6. The performance of mmWave-NOMA and mmWave-OMA with multiple beams. $M = 8$, $N = 4$, $\lambda = 10$, $\Delta = 0.01$, $i = 3$, $j = 1$ and $R_i = 0.5$ BPCU.

sensitive to the choice of the threshold.

In Fig. 6, the performance of the proposed mmWave-NOMA scheme with multiple randomly generated beams is illustrated, where OMA is used as the benchmark again. Different from the previous cases with a single beam, the use of multiple beams means that users in the mmWave-NOMA system suffer more interference. Particularly, even if the strong user in a NOMA pair can use SIC to remove its partner's message, it still experiences interference from the users on other beams. However, the fact that mmWave propagation is highly directional can be used to effectively reduce such inter-beam interference. The reason is that the inter-beam interference, $\sum_{n \neq m} \frac{|a_{m,j}|^2}{(1+d_{m,j}^\alpha)} F_M(\pi[\theta_n - \theta_{m,j}])$, is a function of the angle difference between a user's channel vector and the interference beams. With a choice of $\Delta = 0.01$, i.e., the central angle is about 4 degrees, the inter-beam interference is significantly suppressed, as shown in the two figures. The superior performance of NOMA can also be clearly demonstrated by the fact that the outage probability for the strong user in OMA cannot be reduced to zero, regardless of how large the transmission power is. On the other hand, the use of NOMA can reduce the outage probability rapidly by increasing the transmission power, which is due to the fact that NOMA can realize better spectral efficiency.

Finally, we compare the mmWave-NOMA scheme with

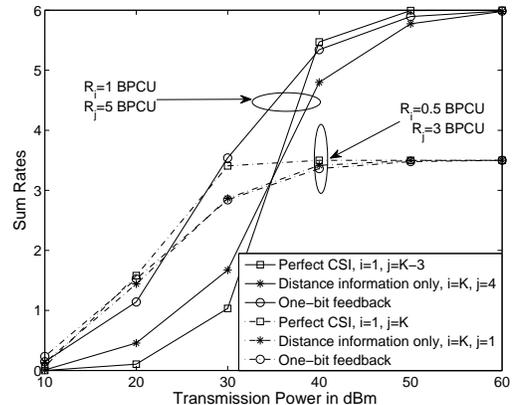


Fig. 7. Sum-rate comparison between the proposed random beamforming transmission schemes. $M = 4$, $K = 5$, $\lambda = 1$, and $\Delta = 0.4$. The threshold is set as $\frac{1}{2}(\bar{\eta}_1 + \bar{\eta}_2)$.

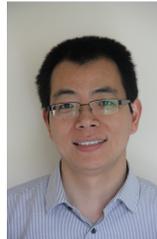
perfect CSI to the two schemes with limited CSI. Intuitively, the cases with limited CSI will result in some performance degradation, but the simulation results in Fig. 7 indicate that the schemes with limited feedback can yield an increase of the system throughput, as explained in the following. Take a four-user case as an example, where the users are ordered as in (8). Suppose that the perfect-CSI based scheme is to schedule user 1 and user 2, i.e., two users with poor channel conditions. Because of the ordering ambiguity caused by the use of partial CSI, the one-bit feedback scheme might schedule user 3 and user 4. According to the broadcast capacity region in [31], scheduling users with better channel conditions yields a larger sum rate, which means that it is possible for the schemes with partial CSI to outperform the one with perfect CSI. Fig. 7 clearly demonstrates this phenomenon. For example, given K users, when the user with the worst channel condition is paired with the user with the second worst channel condition. The schemes with limited feedback can outperform the scheme with perfect CSI, when the transmission power is 20 dBm. It is worth pointing out that a similar observation has been previously reported in [9] in the context of massive MIMO.

VII. CONCLUSIONS

In this paper, we have investigated the coexistence between NOMA and mmWave communications. We have first considered the application of random beamforming to the addressed mmWave-NOMA scenario, by focusing on the case with a single beam generated at the base station. Stochastic geometry has been applied to characterize the performance of the mmWave-NOMA transmission scheme, by using the key features of mmWave networks, i.e., mmWave transmission is highly directional and potential blockages will thin the user distribution. Two beamforming approaches that can effectively reduce feedback have also been proposed to the addressed mmWave-NOMA communication networks, and the performance for the scenario with multiple beams has also been studied. The provided simulation results have demonstrated that the developed analytical results are accurate, and the proposed mmWave-NOMA transmission schemes yield significant performance gains over conventional mmWave-OMA schemes.

REFERENCES

- [1] "5G radio access: Requirements, concepts and technologies," NTT DOCOMO, Inc., Tokyo, Japan, 5G Whitepaper, Jul. 2014.
- [2] "Proposed solutions for new radio access," Mobile and wireless Commun. Enablers for the Twenty-twenty Information Society (METIS), Deliverable D.2.4, Feb. 2015.
- [3] Y. Lan, A. Benjebboiu, X. Chen, A. Li, and H. Jiang, "Considerations on downlink non-orthogonal multiple access (NOMA) combined with closed-loop SU-MIMO," in *Proc. IEEE Int. Conf. on Signal Processing and Commun. Systems*, Gold Coast, Australia, Dec 2014.
- [4] 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.
- [5] 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, Sept. 2015.
- [6] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G non-orthogonal multiple access," *IEEE Trans. Veh. Tech.*, vol. 65, no. 8, pp. 6010–6023, Aug. 2016.
- [7] Z. Ding, L. Dai, and H. V. Poor, "MIMO-NOMA design for small packet transmission in the Internet of Things," *IEEE Access*, vol. 4, pp. 1393–1405, Aug. 2016.
- [8] Y. Liu, Z. Ding, M. ElKashlan, and J. Yuan, "Non-orthogonal multiple access in large-scale underlay cognitive radio networks," *IEEE Trans. Veh. Tech.*, vol. PP, no. 99, pp. 1–1, 2016.
- [9] Z. Ding and H. V. Poor, "Design of massive-MIMO-NOMA with limited feedback," *IEEE Signal Process. Lett.*, vol. 23, no. 5, pp. 629–633, May 2016.
- [10] Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, "Optimal joint power and subcarrier allocation for MC-NOMA systems," in *Proc. IEEE Global Commun. Conf.*, Washington D.C., USA, Dec. 2016.
- [11] T. S. Rappaport, S. Sun, R. Mazyus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.
- [12] M. N. Kulkarni, A. Ghosh, and J. G. Andrews, "A comparison of MIMO techniques in downlink millimeter wave cellular networks with hybrid beamforming," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1952–1967, May 2016.
- [13] G. Lee, Y. Sung, and J. Seo, "Randomly-directional beamforming in millimeter-wave multiuser MISO downlink," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1086–1100, Feb. 2016.
- [14] S. Singh, M. N. Kulkarni, A. Ghosh, and J. G. Andrews, "Tractable model for rate in self-backhauled millimeter wave cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 10, pp. 2196–2211, Oct. 2015.
- [15] J. G. Andrews, R. K. Ganti, M. Haenggi, N. Jindal, and S. Weber, "A primer on spatial modeling and analysis in wireless networks," *IEEE Commun. Mag.*, vol. 48, no. 11, pp. 156–163, Nov. 2010.
- [16] Q. Sun, S. Han, C.-L. I, and Z. Pan, "On the ergodic capacity of MIMO NOMA systems," *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, pp. 405–408, Aug. 2015.
- [17] J. Choi, "On the power allocation for MIMO-NOMA systems with layered transmissions," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3226–3237, May 2016.
- [18] Z. Ding, F. Adachi, and H. V. Poor, "The application of MIMO to non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 537–552, Jan. 2016.
- [19] M. Haenggi, *Stochastic Geometry for Wireless Networks*. Cambridge University Press, Cambridge, UK, 2012.
- [20] T. S. Rappaport, E. Ben-Dor, J. N. Murdock, and Y. Qiao, "38 GHz and 60 GHz angle-dependent propagation for cellular amp; peer-to-peer wireless communications," in *Proc. IEEE Int. Conf. Commun.*, Ottawa, Canada, Jun. 2012.
- [21] S. Biswas, S. Vuppala, J. Xue, and T. Ratnarajah, "On the performance of relay aided millimeter wave networks," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 576–588, Apr. 2016.
- [22] A. Thornburg, T. Bai, and R. Heath, "Performance analysis of mmwave ad hoc networks," *IEEE Trans. Signal Process.*, to appear in 2016.
- [23] G. Lee, Y. Sung, and M. Kountouris, "On the performance of random beamforming in sparse millimeter wave channels," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 560–575, Apr. 2016.
- [24] X. Gao, L. Dai, S. Han, C. L. I, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for mmWave MIMO systems with large antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, Apr. 2016.
- [25] 3rd Generation Partnership Project (3GPP), "Study on downlink multiuser superposition transmission for LTE," Mar. 2015.
- [26] H. Wang, S. Ma, T.-S. Ng, and H. V. Poor, "A general analytical approach for opportunistic cooperative systems with spatially random relays," *IEEE Trans. Wireless Commun.*, vol. 10, no. 12, pp. 4122–4129, Dec. 2011.
- [27] H. A. David and H. N. Nagaraja, *Order Statistics*. John Wiley, New York, 3rd ed., 2003.
- [28] E. Hildebrand, *Introduction to Numerical Analysis*. Dover, New York, USA, 1987.
- [29] L. Zheng and D. N. C. Tse, "Diversity and multiplexing : A fundamental tradeoff in multiple antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1073–1096, May 2003.
- [30] M. Haenggi, "On distances in uniformly random networks," *IEEE Trans. on Information Theory*, vol. 51, no. 10, pp. 3584–3586, Oct. 2005.
- [31] T. Cover and J. Thomas, *Elements of Information Theory*, 6th ed. Wiley and Sons, New York, 1991.



Zhiguo Ding (S'03-M'05) received his B.Eng in Electrical Engineering from the Beijing University of Posts and Telecommunications in 2000, and the Ph.D degree in Electrical Engineering from Imperial College London in 2005. From Jul. 2005 to Aug. 2014, he was working in Queen's University Belfast, Imperial College and Newcastle University. Since Sept. 2014, he has been with Lancaster University as a Chair Professor. From Oct. 2012 to Sept. 2016, he has been also with Princeton University as a Visiting Research Collaborator.

Dr Ding's research interests are 5G networks, game theory, cooperative and energy harvesting networks and statistical signal processing. He is serving as an Editor for *IEEE Transactions on Communications*, *IEEE Transactions on Vehicular Networks*, *IEEE Wireless Communication Letters*, *IEEE Communication Letters*, and *Journal of Wireless Communications and Mobile Computing*. He received the best paper award in IET Comm. Conf. on Wireless, Mobile and Computing, 2009, IEEE Communication Letter Exemplary Reviewer 2012, and the EU Marie Curie Fellowship 2012–2014.



Pingzhi Fan (M93SM99-F15) received his PhD degree in Electronic Engineering from the Hull University, UK. He is currently a professor and director of the institute of mobile communications, Southwest Jiaotong University, China. He is a recipient of the UK ORS Award, the Outstanding Young Scientist Award by NSFC, and the chief scientist of a national 973 research project. He served as general chair or TPC chair of a number of international conferences, and is the guest editor-in-chief, guest editor or editorial member of several international journals. He is the founding chair of IEEE VTS BJ Chapter and ComSoc CD Chapter, the founding chair of IEEE Chengdu Section. He also served as a board member of IEEE Region 10, IET(IEE) Council and IET Asia-Pacific Region. He has over 200 research papers published in various academic English journals (IEEE/IEE/IEICE, etc), and 8 books (incl. edited), and is the inventor of 20 granted patents. His research interests include high mobility wireless communications, 5G technologies, wireless networks for big data, signal design & coding, etc.

He is the founding chair of IEEE VTS BJ Chapter and ComSoc CD Chapter, the founding chair of IEEE Chengdu Section. He also served as a board member of IEEE Region 10, IET(IEE) Council and IET Asia-Pacific Region. He has over 200 research papers published in various academic English journals (IEEE/IEE/IEICE, etc), and 8 books (incl. edited), and is the inventor of 20 granted patents. His research interests include high mobility wireless communications, 5G technologies, wireless networks for big data, signal design & coding, etc.



H. Vincent Poor (S'72, M'77, SM'82, F'87) received the Ph.D. degree in EECS from Princeton University in 1977. From 1977 until 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990 he has been on the faculty at Princeton, where he is currently the Michael Henry Strater University Professor of Electrical Engineering. During 2006 to 2016, he served as Dean of Princetons School of Engineering and Applied Science. His research interests are in the areas of information theory, statistical signal processing and

stochastic analysis, and their applications in wireless networks and related fields such as smart grid and social networks. Among his publications in these areas is the book *Mechanisms and Games for Dynamic Spectrum Allocation* (Cambridge University Press, 2014).

Dr. Poor is a member of the National Academy of Engineering, the National Academy of Sciences, and is a foreign member of the Royal Society. He is also a fellow of the American Academy of Arts and Sciences, the National Academy of Inventors, and other national and international academies. He received a Guggenheim Fellowship in 2002, and the IEEE Education Medal in 2005. Recent recognition of his work includes the 2016 John Fritz Medal, the 2017 IEEE Alexander Graham Bell Medal, Honorary Professorships at Peking University and Tsinghua University, both conferred in 2016, and a D.Sc. *honoris causa* from Syracuse University awarded in 2017.