On the Coexistence of Non-orthogonal Multiple Access and Millimeter-Wave Communications

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

***Abstract*—In this paper, the application of non-orthogonal multiple access (NOMA) to millimeter-wave (mmWave) communications is considered, where random beamforming is used in order to reduce the system overhead. 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. The provided analytical and numerical results demonstrate that the combination of NOMA and mmWave yields significant gains in terms of sum rates and outage probabilities, compared to conventional orthogonal multiple access based mmWave systems.**

I. Introduction

Non-orthogonal multiple access (NOMA) and millimeter-wave (mmWave) communications have been recognized as two key enabling technologies in the fifth generation (5G) mobile networks [1]–[4]. Both the technologies are motivated by the same phenomenon, the spectrum crunch, but the solutions provided by them are different. NOMA is to use the available spectrum more efficiently, and mmWave transmission is to use mmWave bands which are less occupied compared to those used by current cellular networks. Even though more bandwidth resources are available in very high frequencies, the use of NOMA is still important since the huge demands on bandwidth resources due to the exponential growth of broadband traffic can be met only by acquiring more radio spectrum and also efficiently using these acquired spectrum.

In this paper, we consider an mmWave-NOMA downlink scenario where a base station equipped with multiple antennas communicates with multiple single-antenna nodes. While MIMO-NOMA has been extensively studied in [5]–[7], 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 described in the following. We first consider the application of random beamforming to the addressed mmWave-NOMA scenario, where 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 a lot of 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 further reduce 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 is also used to model the user distribution more realistically.

In a fast time varying situation, where 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 loss, instead of their effective channel gains. The impact of this partial channel state information (CSI) on the performance of the mmWave-NOMA networks is investigated.

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 the 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. The exact expressions for the outage probabilities achieved by the random beamforming scheme are developed to characterize the performance of mmWave-NOMA communications.

II. System Model

Consider an mmWave-NOMA downlink transmission scenario with one base station communicating with multiple users. The base station is equipped with M antennas and each user has a single antenna. Denote the disc which is covered by the base station by $\mathcal{D}$. Assume that the base station is located at the origin of $\mathcal{D}$ and the radius of the disk is denoted by $R_{\mathcal{D}}$. Assume that users are randomly deployed in the disc following the homogeneous Poisson point process (HPPP), with the density of λ [8]. Therefore, the number of the users in the disc is Poisson distributed, i.e., $P(\text{K users in } \mathcal{D}) = \frac{\mu^K e^{-\mu}}{K!}$, where $\mu = \pi R_{\mathcal{D}}^2\lambda$.

As discussed in [4] and [9], the mmWave channel model is quite different from those of conventional lower frequency cellular networks, and 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, θ_k^l denotes the normalized angle-of-departure of the l-th path,

$$\mathbf{a}(\theta) = \frac{1}{\sqrt{M}}\begin{bmatrix} 1 & e^{-j\pi\theta} & \cdots & e^{-j\pi(M-1)\theta}\end{bmatrix}^T, \quad (2)$$

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, $a_{k,l}$

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.

denotes the complex gain for the l-th path and is complex Gaussian distributed, i.e., $a_{k,l} \sim CN(0,1)$. As discussed in [9] and [10], the path loss of NLOS components will be much more significant than that of the LOS component, and therefore the first factor at the right-hand side of (1) is dominant, which yields the following simplified channel model

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

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

In practice, the direct path between the transceivers might be blocked by obstacles, which means that LOS does not always exist. As a result, in addition to path loss and fading attenuation, mm-Wave transmission also suffers potential blockages, which is an important feature to be captured. As shown in [11] and [12], this blockage feature can be modelled as follows:

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

where ϕ is determined by the building density, the shape of the buildings, etc. 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 huge amount of system overhead, if there are many users in the system. In order to reduce the system overhead, this paper is to 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 that a single beam, denoted by $\mathbf{p}$, is generated at the base station. Since analog precoding is more preferable to mmWave systems, we use the following choice for beamforming

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

where $\bar{\theta}$ is uniformly distributed between -1 and 1.

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 a user which has the strongest channel condition. However, such an approach will still consume a lot of system overhead, particularly if there are a lot of users in the cell.

In the context of mmWave communications, a useful observation is that many users do not have to participate into the competition for the access to the channel, as explained in the following. Without loss of generality, user m 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:

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

Following steps similar to those in [9], this effective channel gain can be rewritten as follows:

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

where $F_M(x)$ denotes the Fejér kernel. Note that a Fejér kernel goes to zero quickly by increasing its parameter, i.e., $F_M(x)\rightarrow 0$ by increasing x. This means that a user can have a large effective channel gain on the 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 a circular sector. Particularly, this circular sector is denoted by $\mathcal{D}_\theta$, 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.

B. The Implementation of NOMA

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

$$|\mathbf{h}_1^H\mathbf{p}|^2 \leq \cdots \leq |\mathbf{h}_K^H\mathbf{p}|^2. \tag{8}$$

Similarly to [13] and [5], we consider the case that two users will be selected for the implementation of NOMA. Note that the implementation of NOMA in long term evolution advanced (LTE-A) is aso based on the two-user case [14]. Without loss of generality, we consider user i and user j, $1\leq i<j\leq K$, are paired together for NOMA transmission on the randomly generated beam, which means the signal sent by the base station is given by $\mathbf{p}\left(\beta_i s_i+\beta_j s_j\right)$, 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

$$y_i = \mathbf{h}_i^H\mathbf{p}\left(\beta_i s_i+\beta_j s_j\right)+n_i, \tag{9}$$

where n_i denotes the 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)

$$\mathrm{SINR}_i = \frac{|\mathbf{h}_i^H\mathbf{p}|^2\beta_i^2}{|\mathbf{h}_i^H\mathbf{p}|^2\beta_j^2+\frac{1}{\rho}}, \tag{10}$$

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

$$\mathrm{P}_{i|K}^o = \mathrm{P}\left(\mathrm{SINR}_i<\epsilon_i|K\right), \tag{11}$$

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

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

$$\mathrm{SINR}_j = \rho|\mathbf{h}_j^H\mathbf{p}|^2\beta_j^2, \tag{12}$$

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

$$\mathrm{P}^o_{j|K} = 1 - \mathrm{P}\left(\mathrm{SINR}_{i\to j} > \epsilon_i, \mathrm{SINR}_j > \epsilon_j | K\right), \tag{13}$$

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:

$$R^{NOMA}_{sum} = \mathrm{P}(K=1)(1-\mathrm{P}^{1|K}_{OMA})R_1 + \sum_{k=2}^{\infty} \mathrm{P}(K=k) \times \left((1-\mathrm{P}^o_{i|K})R_i + (1-\mathrm{P}^o_{j|K})R_j\right), \tag{14}$$

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

$$R^{OMA}_{sum} = \mathrm{P}(K=1)(1-\mathrm{P}^{1|K}_{OMA})R_1 + \sum_{k=2}^{\infty} \mathrm{P}(K=k) \times \left((1-\mathrm{P}^{i|K}_{OMA})R_i + (1-\mathrm{P}^{j|K}_{OMA})R_j\right), \tag{15}$$

where $\mathrm{P}^{n|K}_{OMA}$ denotes the conditional outage probability when OMA is used. The reason for using the OMA mode in (14) is that it might be possible that there is only one user in $\mathcal{D}_\theta$. In this case, NOMA cannot be implemented and we simply use OMA, i.e., $\mathrm{P}^{n|K}_{OMA} = \mathrm{P}\left(\log(1+\rho|\mathbf{h}_n^H\mathbf{p}|^2) < 2R_n\right)$, for $n \in \{i, j\}$.

C. Characterize the Sum Rate and Outage Probabilities

In order to evaluate the sum rate shown in (14), it is important to find the density function of the ordered channel gain, $|\mathbf{h}_j^H\mathbf{p}|^2$, which can be accomplished by first characterizing the unordered channel gains and then applying 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 probability density function (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:

$$F_{\pi(j)}(y) = \int_{\mathcal{D}_\theta} \mathrm{P}\left(|\mathbf{h}_{\pi(j)}^H\mathbf{p}|^2 < y \mid X_{\pi(j)} = x_{\pi(j)}\right) dP_{X_{\pi(j)}} = \int_{\mathcal{D}_\theta} \left(1 - e^{-\frac{y(1+r(x)^\alpha)}{F_M\left(\pi[\bar{\theta}-\theta_{\pi(j)}]\right)}}\right) p_{X_{\pi(j)}}(x)dx, \tag{16}$$

where $r(x)$ denotes the distance from the origin to point x. Note that the condition of 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 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 a LOS path. Therefore, following the discussions in [15], the effect of blockages is to thin the original homogeneous point process and this thinning process yields another PPP with the following intensity

$$\lambda_{\Phi_2}(x) = \lambda e^{-\phi r(x)}. \tag{17}$$

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

$$\mu_{\Phi_2}(\mathcal{D}_\theta) = 2\Delta\lambda\phi^{-2}\gamma(2, R_{\mathcal{D}}\phi), \tag{18}$$

where $\gamma(\cdot)$ denotes the incomplete gamma function. As a result, after considering potential blockages, the probability to have K users in the sector, $\mathcal{D}_\theta$, can be obtained as follows:

$$\mathrm{P}(K=k) = \frac{\left(\mu_{\Phi_2}(\mathcal{D}_\theta)\right)^k}{k!} e^{-\mu_{\Phi_2}(\mathcal{D}_\theta)}. \tag{19}$$

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

$$p_{X_{\pi(j)}}(x) = \frac{\lambda_{\Phi_2}(x)}{\mu_{\Phi_2}(\mathcal{D}_\theta)} = \frac{\lambda\phi^2 e^{-\phi r(x)}}{2\Delta\lambda\gamma(2, R_{\mathcal{D}}\phi)}. \tag{20}$$

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

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

where the last equation follows by using polar coordinates. After using the fact that all the channel gains are independent and identically distributed and also applying order statistics, we can find that the pdf of the ordered channel gain is given by [16]

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

where $c_j = \frac{K!}{(j-1)!(K-j)!}$. Therefore the outage probability experienced by user j conditioned on K can be calculated as follows:

$$\mathrm{P}^o_{j|K} = 1 - \mathrm{P}\left(\mathrm{SINR}_{i\to j} > \epsilon_i, \mathrm{SINR}_j > \epsilon_j | K\right) = c_j \sum_{p=0}^{K-j} \binom{K-j}{p} (-1)^p \frac{F^{j+p}_{\pi(j)}(\eta_j)}{j+p}, \tag{21}$$

if $\beta_i^2 > \beta_j^2\epsilon_i$, otherwise $\mathrm{P}^o_{j|K} = 1$, where $\eta_j = \max\left\{\frac{\frac{\epsilon_i}{\rho}}{\beta_i^2 - \beta_j^2\epsilon_i}, \frac{\epsilon_j}{\rho\beta_j^2}\right\}$.

Similarly the conditional outage probability for user i can be obtained as follows:

$$\mathrm{P}^o_{i|K} = c_i \sum_{p=0}^{K-i} \binom{K-i}{p} (-1)^p \frac{F^{i+p}_{\pi(j)}(\eta_i)}{i+p}, \tag{22}$$

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

By substituting (19), (21) and (22) into (14), the sum rate achieved by mmWave NOMA can be calculated. By using the fact that mmWave transmission is highly directional, one can develop more insightful approximations for the sum rates and the outage probabilities as shown in [17], which will be omitted here due to the space limits.

IV. Random Beamforming with Partial CSI

In the previous section, it is assumed that the base station has the perfect knowledge of the users' effective channel gains, which might not be realistic in a fast time varying situation. Compared to the phases and fading of the channels, the distance information of the users is changing relatively slowly. Therefore, in this section, we investigate the impact of this partial CSI on the performance of mmWave NOMA.

Again consider that only the users which fall into the sector $\mathcal{D}_\theta$ will participate into the NOMA transmission. Assume that there are K users in this sector. Since the user's distances are known, the base station will order the users according to the following criterion

$$d_1 \leq \cdots \leq d_K, \tag{23}$$

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 NOMA transmission to act as the weak and strong users, respectively. Since a user with a shorter distance has a stronger channel condition, we let $i > j$.

Note that the density functions of the ordered distances have been found in [18] by considering a ball area. The shape of the addressed area is a sector, but the steps provided in [18] are still applicable, as shown in the following. Particularly, the CDF of d_k can be calculated from the probability that there are less than k users inside the sector with radius of r, i.e.,

$$F_{d_k}(r) = 1 - \sum_{i=0}^{k-1} \mathrm{P}(E_i) = 1 - \sum_{i=0}^{k-1} e^{-\mu_{\Phi_2}(\mathcal{A}(r))} \frac{\left(\mu_{\Phi_2}(\mathcal{A}(r))^i\right)}{i!},$$

where E_i denotes the event that there are i users in the sector with radius of r and $\mathcal{A}(r)$ denotes a sector with radius of r. Following steps similar to those for obtaining (18), the factor $\mu_{\Phi_2}(\mathcal{A}(r))$ can be found as follows:

$$\mu_{\Phi_2}(\mathcal{A}(r)) = 2\Delta\lambda\phi^{-2}\gamma(2, r\phi). \tag{24}$$

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

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

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

$$f_{d_k}(r) = 2\Delta\lambda e^{-r\phi} r e^{-2\Delta\lambda\phi^{-2}\gamma(2,r\phi)} \frac{\left(2\Delta\lambda\phi^{-2}\gamma(2, r\phi)\right)^{k-1}}{(k-1)!}, \tag{25}$$

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 [18] 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 angel of user k's channel vector is independent from its distance, and it is uniformly distributed between $(\bar{\theta} - \Delta)$ and $(\bar{\theta} + \Delta)$. Therefore the CDF of user k's channel gain can be obtained as follows:

$$\begin{aligned} F_k(y) &= \int\limits_{\mathcal{D}_\theta} \left(1 - e^{-\frac{y(1+r(x)^\alpha)}{F_M\left(\pi[\bar{\theta}-\theta_{\pi(j)}]\right)}}\right) p_{X_{\pi(j)}}(x) dx \\ &= \int_{\bar{\theta}-\Delta}^{\bar{\theta}+\Delta} \int_0^{R_\mathcal{D}} \left(1 - e^{-\frac{y(1+r^\alpha)}{F_M\left(\pi[\bar{\theta}-\theta]\right)}}\right) \frac{f_{d_k}(r)}{2\Delta} dr d\theta. \end{aligned} \tag{26}$$

It is important to point out that the above CDF is valid only if we can find the k-th nearest node. On in other words, if there is no boundary of $\mathcal{D}_\theta$ and the nodes are spreading throughout of the plane, the above CDF can be applied. For the addressed scenario, the users are confined in $\mathcal{D}_\theta$, i.e., $r \leq R_\mathcal{D}$, which means that it might be possible that the k-th nearest node does not exist, i.e., there are less than $(k-1)$ nodes in $\mathcal{D}_\theta$. As a result, the outage probability for the k-th nearest node can be written as follows: [1]

$$F_k^o = \sum_{n=0}^{k-1} \mathrm{P}(K = n) + \left(1 - \sum_{n=0}^{k-1} \mathrm{P}(K = n)\right) F_k(\eta_k), \tag{27}$$

where $k \in \{i, j\}$. The factor $\sum_{n=0}^{k-1} \mathrm{P}(K = n)$ denotes the probability for the events that the k-th nearest node cannot be found in $\mathcal{D}_\theta$.

Therefore, the outate sum rate can be written as follows:

$$R_{sum}^{NOMA} = (1 - F_j(\eta_j))R_j + (1 - F_i(\eta_i))R_i, \tag{28}$$

where the i-th nearest user has the targeted data rate of R_i. While the above expressions about the outage probability and the sum rate can be calculated numerically, more simplified approximations can be obtained, as shown in [17].

V. Random Beamforming: A Multiple-Beam Case

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 the transmission. Following [9], these N orthonormal beamforming vectors can be constructed as follows:

$$\mathbf{p}_m = \mathbf{a}\left(\zeta + \frac{2(m-1)}{N}\right), \tag{29}$$

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 $\bar{\theta}_m$.

Priori to downlink transmission, the base station will first broadcast pilot signals on these K orthogonal beams. Similarly to $\mathcal{D}_\theta$, define $\mathcal{D}_{\theta_m}$ as the circular sector around $\bar{\theta}_m$ with a central angle of 2Δ. Only the users which fall into the sector $\mathcal{D}_{\theta_m}$ will participate into the NOMA transmission on beam m. Particularly, 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

[1] One can also use a CDF expression conditioned on K, but this is difficult to evaluate since the condition of K converts the Poisson point process to a Bernoulli one to which the result in (25) is not applicable.

i and user j on beam m, to act as the weak and strong users, respectively.

Therefore, the base station will superimpose the two users' messages on the m-th beam as follows:

$$\sum_{m=1}^{N} \mathbf{p}_m \left(\beta_{m,1} s_{m,i} + \beta_{m,2} s_{m,j}\right), \tag{30}$$

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

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

$$\begin{aligned} y_{m,j} =& \mathbf{h}_{m,j}^H \mathbf{p}_m \left(\beta_{m,1} s_{m,i} + \beta_{m,2} s_{m,j}\right) \\ &+ \mathbf{h}_{m,j}^H \sum_{n=1, n\neq m}^{K_m} \mathbf{p}_n \left(\beta_{n,1} s_{n,i} + \beta_{n,2} s_{n,j}\right) + n_{m,j}, \end{aligned} \tag{31}$$

where $n_{\mathcal{S}_1,j}$ denotes the 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:

$$\text{SINR}_{m,i\rightarrow j} = \frac{|\mathbf{h}_{m,j}^H \mathbf{p}_m|^2 \beta_{m,1}^2}{|\mathbf{h}_{m,j}^H \mathbf{p}_m|^2 \beta_{m,2}^2 + \sum\limits_{n\neq m} |\mathbf{h}_{m,j}^H \mathbf{p}_n|^2 + \frac{1}{\rho}}. \tag{32}$$

If $\text{SINR}_{m,i\rightarrow j} \geq \epsilon_{m,i}$, intra-group interference can be cancelled and the user can decode its own information with the following SINR

$$\text{SINR}_{m,j} = \frac{|\mathbf{h}_{m,j}^H \mathbf{p}_m|^2 \beta_{m,2}^2}{\sum\limits_{n\neq m} |\mathbf{h}_{m,j}^H \mathbf{p}_n|^2 + \frac{1}{\rho}}. \tag{33}$$

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

$$\text{SINR}_{m,i} = \frac{|\mathbf{h}_{m,i}^H \mathbf{p}_m|^2 \beta_{m,1}^2}{|\mathbf{h}_{m,i}^H \mathbf{p}_m|^2 \beta_{m,2}^2 + \sum\limits_{n\neq m} |\mathbf{h}_{m,i}^H \mathbf{p}_n|^2 + \frac{1}{\rho}}. \tag{34}$$

Different to 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 is depending on two different SINR functions, $\text{SINR}_{m,i\rightarrow j}$ and $\text{SINR}_{m,j}$. For the illustration purpose, we focus on a simple user scheduling scheme based on distances, a strategy similar to the one proposed in Section IV. Therefore we can order these users who will participate in the NOMA transmission on beam m as follows:

$$d_{m,1} \leq \cdots \leq d_{m,K_m}, \tag{35}$$

where K_m is the number of users in the sector $\mathcal{D}_{\theta_m}$. Furthermore suppose that user i has a distance larger than that of user j, i.e., $i > j$.

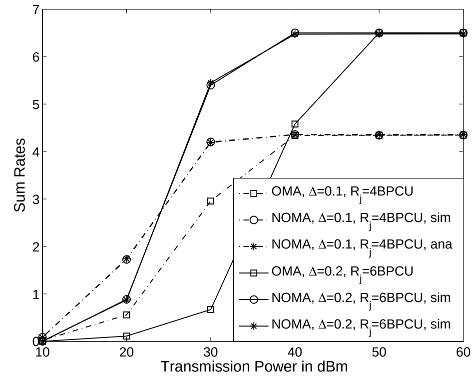

Fig. 1. The performance of mmWave-NOMA and mmWave-OMA with perfect CSI. $M = 4$, $\lambda = 1$, $\Delta = 0.1$, $R_i = 0.5$ BPCU, $i = 1$ and $j = K$.

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

$$\mathrm{P}_{m,j}^o = 1 - \mathrm{P}\left(\text{SINR}_{m,i\rightarrow j} > \epsilon_{m,1}, \text{SINR}_{m,j} > \epsilon_{m,2}\right), \tag{36}$$

where $\epsilon_{m,1} = 2^{R_{m,1}} - 1$, $R_{m,1}$ denotes the targeted rate for user i on beam m, $\epsilon_{m,2}$ and $R_{m,2}$ are defined similarly.

Unlike those SINR functions in the previous sections, the SINRs for the case with multiple beams are more complicated. With some algebraic manipulations shown in [17], the outage probability of user j on beam m can be expressed as follows:

$$\begin{aligned} \mathrm{P}_{m,j}^o =& 1 - \int_{\bar{\theta}-\Delta}^{\bar{\theta}+\Delta} \int_0^{R_{\mathcal{D}}} \exp\{-\max\{ \\ & \frac{\frac{\epsilon_{m,1}}{\rho}(1 + d_{m,j}^\alpha)}{F_{j,m}^m \beta_{m,1}^2 - \epsilon_{m,1} F_{j,m}^m \beta_{m,2}^2 - \sum\limits_{n\neq m} \epsilon_{m,1} F_{j,n}^m}, \\ & \frac{\frac{\epsilon_{m,2}}{\rho}(1 + d_{m,j}^\alpha)}{F_{j,m}^m \beta_{m,2}^2 - \sum\limits_{n\neq m} \epsilon_{m,2} F_{j,n}^m}\}\} \frac{f_{d_j}(r)}{2\Delta} dr d\theta, \end{aligned} \tag{37}$$

if $F_{j,m}^m \beta_{m,1}^2 > \epsilon_{m,1} F_{j,m}^m \beta_{m,2}^2 + \sum\limits_{n\neq m} \epsilon_{m,1} F_{j,n}^m$ and $F_{j,m}^m \beta_{m,2}^2 > \sum\limits_{n\neq m} \epsilon_{m,2} F_{j,n}^m$, otherwise the outage probability will be always one, where $F_{j,n}^m \triangleq F_M\left(\pi[\bar{\theta}_n - \theta_{m,j}]\right)$. The outage probability for user i can be obtained similarly. Following those steps in [17], one can further simplify the above analytical results, which will be omitted here due to space limits.

VI. Numerical Studies

In this section, the performance of the proposed mmWave-NOMA transmission schemes are evaluated by using computer simulations. The path loss exponent is set as $\alpha = 2$, the radius of $\mathcal{D}$ is $R_{\mathcal{D}} = 10$m, the noise power is -30dBm, the blockage parameter is set as $\phi = 0.1$, $\beta_i^2 = \frac{3}{4}$ and $\beta_j^2 = \frac{1}{4}$ are used as the NOMA power allocation coefficients.

In Fig. 1, the performance of the proposed random beamforming scheme in mmWave-NOMA systems is studied, where the mmWave-OMA scheme is used as a benchmark. As can be observed from Fig.1, the use of NOMA can yield a significant sum rate gain over the OMA scheme, and this gain can be enlarged when the targeted data rate of the strong user is increased. For example, for $R_j = 4$ bit per channel user (BPCU), the gain of mmWave-NOMA over mmWave-OMA is

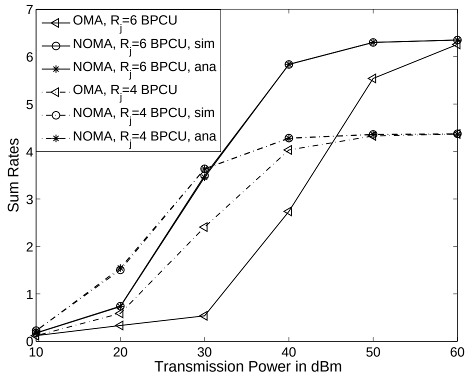

Fig. 2. The performance of mmWave-NOMA and mmWave-OMA by using the distance information only. $M = 4$, $\lambda = 1$, $\Delta = 0.1$, $R_i = 0.5$ BPCU, $i = 4$ and $j = 1$.

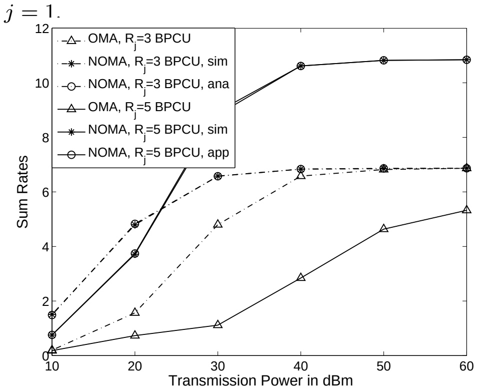

Fig. 3. The performance of mmWave-NOMA and mmWave-OMA with multiple beams. $M = 8$, $N = 4$, $\lambda = 10$, $\Delta = 0.01$, $i = 3$, $j = 1$ and $R_i = 0.5$.

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. In Fig. 2, the performance of the mmWave-NOMA and mmWave-OMA schemes is compared, where the base station has access to the users' distance information only. As can be observed from the figure, the use of NOMA can yield a significant performance gain in the sum rate, compared to the OMA scheme, even if only the distance information is available to the base station. It is also important to point out that the developed exact expressions for the sum rate match the simulation results perfectly.

In Fig. 3, the performance of the proposed mmWave-NOMA scheme with multiple randomly generated beams is demonstrated, where the OMA scheme is used as the benchmarking scheme again. Different to 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 parter's message, it still suffers the interference from the users on other beams. However, the feature of mmWave networks that mmWave prorogation is highly directional can effectively reduce such inter-beam interference. The reason is that the inter-beam interference, $\sum\limits_{n \neq m} \frac{|a_{m,j}|^2}{(1+d_{m,j}^{\alpha})} F_M\left(\pi[\bar{\theta}_n - \theta_{m,j}]\right)$, 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 angel is about 4 degrees, the inter-beam interference can be significantly suppressed, as shown in the figure.

VII. Conclusions

In this paper, we have investigated the coexistence between NOMA and mmWave communications, by using random beamforming as a case study. Stochastic geometry has been applied to characterize the performance of the mmWave-NOMA transmission scheme. The provided analytical and simulation results have demonstrated the superior performance of the proposed mmWave-NOMA transmission schemes.

References

[1] 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 International Conference on Signal Processing and Communication Systems*, Gold Coast, Australia, Dec 2014, pp. 1–5.

[2] 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.

[3] T. S. Rappaport, S. Sun, R. Mayzus, 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.

[4] 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.

[5] 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.

[6] J. Choi, "On the power allocation for MIMO-NOMA systems with layered transmissions," *IEEE Trans. Wireless Commun.*, (to appear in 2016).

[7] 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.

[8] M. Haenggi, *Stochastic Geometry for Wireless Networks*. Cambridge University Press, Cambridge, UK, 2012.

[9] 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.

[10] 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 International Conference on Communications*, Ottawa, Canada, Jun. 2012.

[11] 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. Magazine*, vol. 48, no. 11, pp. 156–163, Nov. 2010.

[12] A. Thornburg, T. Bai, and R. Heath, "Performance analysis of mmwave ad hoc networks," *IEEE Trans. Signal Process.*, to appear in 2016.

[13] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G non-orthogonal multiple access," *IEEE Trans. Veh. Tech.*, (to appear in 2016) Available on-line at arXiv:1412.2799.

[14] 3rd Generation Partnership Project (3GPP), "Study on downlink multiuser superposition transmission for LTE," Mar. 2015.

[15] 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.

[16] H. A. David and H. N. Nagaraja, *Order Statistics*. John Wiley, New York, 3rd ed., 2003.

[17] Z. Ding, P. Fan, and H. V. Poor, "Random beamforming in millimeter-wave NOMA networks," *IEEE Trans. Commun.*, (submitted) Available on-line at arxiv.org/abs/1607.06302.

[18] M. Haenggi, "On distances in uniformly random networks," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3584–3586, Oct. 2005.