Performance Analysis of Agile-Beam NOMA in Millimeter Wave Networks

JIANFENG LI, XIAOJUN JING, YANGYING ZHANG, AND JUNSHENG MU, (Student Member, IEEE)
School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China
Corresponding author: Junsheng Mu (mu_junsheng@126.com)

ABSTRACT We propose an agile-beam non-orthogonal multiple access (NOMA) scheme for millimeter wave (mmWave) communication networks. The agile-beam NOMA scheme flexibly switches between the single/multi-beam NOMA depending on the user pairing results. In particular, a novel user pairing strategy employing both distance and angle information is designed to facilitate the agile-beam NOMA transmission. More specifically, the base station first selects one user from the far user group randomly and then pairs it with the user that has the minimum angle difference in the near user group. This pairing strategy efficiently exploits the channel sparsity in user domain while supporting flexible mmWave transmission with agile-beam NOMA. Then an in-depth performance analysis is carried out to reveal the interplay between the key system parameters and the coverage probability of the proposed agile-beam NOMA by using the tool of stochastic geometry. Moreover, numerical results demonstrate the superiority of the proposed agile-beam NOMA over the conventional NOMA and orthogonal multiple access (OMA) in mmWave communication networks.

INDEX TERMS Millimeter wave (mmWave) communication, non-orthogonal multiple access (NOMA), user pairing.

I. INTRODUCTION
Given the spectrum congestion in sub-6 GHz communication networks, millimeter wave (mmWave) communication has been considered as a promising enabling technology in 5G and beyond wireless networks with a much higher capacity to meet explosive demands of traffic [1], [2]. When applying mmWave communication to wireless networks, the benefit of mmWave wireless communication will highly rely on the multiple access strategies, which support multi-user communications simultaneously. As a promising multiple access technique, non-orthogonal multiple access (NOMA) has received considerable attention due to its massive connectivity and superior spectrum efficiency. Different from conventional orthogonal multiple access (OMA) techniques, e.g., TDMA/CDMA/FDMA, NOMA is able to multiplex multiple users in the same (time/frequency/code) resource block with different power levels, i.e., power-domain NOMA, and achieve a higher spectral efficiency by superimposed transmission at transmitter and successive interference cancellation (SIC) at receiver [3]–[6]. This paper focuses on exploring NOMA in mmWave systems.

Applying NOMA in mmWave communication networks has recently received considerable attention and has been investigated in [7]–[13]. Considering the directional transmission with severe propagation path loss in the mmWave band, a large antenna array is equipped at BS to perform directional beamforming for compensating the large path loss. Recently, both single-beam beamforming [7]–[9] and multi-beam beamforming [10]–[13] are employed for NOMA transmission in mmWave networks. A random beamforming NOMA, i.e., single-beam NOMA transmission, is proposed in [7] and outage performance analysis was investigated. In the random beamforming NOMA scheme, the direction of beam is randomly generated by BS with a single radio frequency (RF) chain, and multiple users located within the beamwidth have similar angle-of-departure (AOD) and they are served simultaneously with NOMA scheme. In general, the single-beam NOMA scheme is applicable to the scenario where users are densely deployed and it is always possible to find multiple users located within the beamwidth. However, the beamwidth is such narrow in mmWave communication.
networks that the multiplexed users having separated AODs are not within the same beam direction. Therefore, when users have separated AODs, the single-beam NOMA scheme is no longer applicable, and a multi-beam NOMA scheme [12], [13] was proposed in mmWave communication networks. In the multi-beam NOMA scheme, a beam splitting technique is employed to generate multiple beams and each beam steers directly to one of paired users. The multi-beam NOMA scheme exploits the channel sparsity in the user domain and is thus applicable to the scenario where users are deployed with separated AODs. However, when users are not deployed very densely or very sparsely, which is often the case, this paper attempts to introduce a more flexible mmWave NOMA scheme applicable to the often case.

On the other hand, user pairing is one of the major techniques in NOMA systems and has been widely studied in [14]–[16]. With perfect channel state information (CSI), an asymptotically optimal joint user scheduling and power allocation algorithm based on matching theory was proposed to maximize the sum rate in [17]. Based on monotonic optimization, an effective power allocation and user scheduling algorithm was proposed in [18]. Furthermore, user pairing is investigated with the partial CSI rather than perfect CSI, i.e., channel gain, distance information and angle information. The impact of user pairing on NOMA downlink transmission was investigated in [19], [20], which conclude that the users with more distinctive channel conditions are preferred to be selected as a group to enlarge the performance of NOMA. Two user pairing algorithms taking the users’ channel gain into account are proposed to improve the capacity and guarantee fairness [21]. To further reduce the overhead of channel estimation, BS pairs users according to the distance information instead of channel gain, which is named as distance-based user pairing strategy. Three user pairing strategies based on distance information, i.e., random near user and random far user (RNRF), the nearest near user and the nearest far user (NNNF), as well as the nearest near user and the farthest far user (NNFF), are proposed for simultaneous wireless information and power transfer in NOMA networks [22]. In mmWave NOMA networks, some literature follow the distance-based user pairing strategy, e.g., [7], [8], [23]. However, given narrow beam in mmWave, the users is shown to be more distinctive with respect to their angle information than distance information. Therefore, it is worth pointing out that the angle information has a significant potential in user pairing because of the sparsity in angle domain and directionality of mmWave channel [24]. Recently, an angle-based user pairing strategy is proposed for mmWave NOMA networks [25]. As a result, both the angle information and distance information play a great role in user pairing in mmWave NOMA, and the integration of angle and distance information for user pairing is required.

In this paper, we propose an agile-beam NOMA scheme in mmWave downlink transmission, and then investigate the coverage probability and sum rate of the proposed NOMA scheme. More specifically, we first focus on a typical and representative scenario where a BS equipped with a uniform linear array (ULA) communicates with two NOMA users selected by our proposed user pairing strategy that considers both of the angle and distance information. Subsequently, BS switches flexibly between the single-beam and multi-beam NOMA transmission depending on the angle difference of the paired users. Based on the proposed agile-beam NOMA scheme and user pairing strategy, an in-depth performance analysis is conducted to reveal the interplay between the key system parameters and coverage probability by using the tool of stochastic geometry. In summary, Table I, at the top of next page, shows the comparison between this article and mmWave NOMA literature aforementioned, and the key contributions of this paper are listed as follows.

- An agile-beam NOMA scheme in conjunction with a novel user pairing strategy is proposed for mmWave communication networks, which is applicable to a wide range of scenarios with an adaptive capability to flexibly switch between the single-beam or multi-beam NOMA transmission. To enable such agile-beam NOMA transmission, in contrast to the existing distance-based and angle-based user pairing strategies, we integrate the distance and angle information of users and propose a user pairing strategy, which takes the channel sparsity into account.

### Table 1. A comparison to mmWave NOMA articles.

| Ref | UE density requirement | Beam | Beamforming design | User pairing |
|-----|------------------------|------|---------------------|-------------|
|     |                        |      |                     | Perfect CSI | Channel gain | Angle information | Distance information |
| [7] | Dense                  | Single | ✓                   |             |             | ✓                | ✓                |
| [8] | Dense                  | Single | ✓                   |             |             | ✓                | ✓                |
| [9] | Dense                  | Single | ✓                   |             |             | ✓                | ✓                |
| [10] | No                    | Multiple | ✓                 |             |             |                   | Not mentioned |
| [13] | No                    | Multiple | ✓                 |             |             | ✓                | ✓                |
| [25] | No                    | Single | ×                   |             |             | ✓                | ✓                |
| This paper | Flexible | Single or multiple | ✓ |                     |             | ✓                | ✓                |

VOLUME 8, 2020
of angle-domain in mmWave band into account. The proposed agile-beam NOMA scheme overcomes the constraint of user distribution and ensures BS to serve the paired users with NOMA transmission.

- An in-depth performance analysis is carried out to evaluate the benefits of the proposed agile-beam NOMA scheme in mmWave networks. By utilizing the tool of stochastic geometry, the theoretical coverage probability is derived. Aiming to obtain some insights, asymptotic analysis is carried out, which offers concise and accurate approximations that reveals the interplay between key system parameters and the system utility. Then, the sum rate of the proposed agile-beam NOMA scheme in mmWave networks is obtained with the approximate results of coverage probability.

- Furthermore, we compare the performance of the proposed agile-beam NOMA scheme with the conventional NOMA scheme and OMA scheme in terms of coverage probability and sum rate. Simulation results validate the performance analysis and demonstrate the superiority of the proposed agile-beam NOMA scheme.

The remainder of this paper is organized as follows. We first describe the networks topology, user pairing strategy and channel model in Section II. Then, the coverage probability and sum rate are derived, and the asymptotic analysis is carried out in Section III. The extension of the proposed agile-beam NOMA scheme to multiple-RF chain system is provided in Section IV. Simulation results are provided in Section V. Lastly, Section VI concludes this paper.

Symbol Notation: Hereafter, \( j = \sqrt{-1} \), \( a \) and \( a \) designate a scalar and a vector, respectively. Superscripts \((\cdot)^T\) and \((\cdot)^H\) denote transpose and conjugate transpose operation, respectively. \(|\cdot|\) denotes the absolute value of a complex scalar.

II. SYSTEM MODEL

In this section, we first present the network topology in the considered agile-beam mmWave NOMA system followed by a detailed description of our proposed user pairing strategy and flexible switch between the single-beam and multi-beam NOMA transmission based on the user pairing results. Then, some assumptions and channel model are provided.

A. NETWORK TOPOLOGY

Without loss of generality, we consider a downlink transmission scenario in agile-beam mmWave NOMA network with a BS and two groups of spatially random users, i.e., \( \{A_i\} \) and \( \{B_i\} \), equipped with a single antenna, as shown in Fig.1. The BS is located at the center, and it is equipped with a ULA of \( M \) antenna elements, and the inter-antenna space is a half wavelength. We assume that the near users in group \( \{A_i\} \) are deployed within a disc with radius \( R_1 \). The far users in group \( \{B_i\} \) are deployed within a disc with inner and outer radius \( R_2 \) and \( R_3 \) (assuming \( R_2 < R_1 \) so that the channel conditions are different for the two groups of users. Meanwhile, the users in the ring with radius of \( R_1 \) and \( R_2 \) are included.) The near users and far users are deployed according to independent homogeneous Poisson Point Processes (PPPs) \( \Phi_1 \) and \( \Phi_2 \) with densities \( \lambda_1 \) and \( \lambda_2 \), respectively.

B. USER PAIRING STRATEGY

The decoding complexity of SIC at receiver terminal increases with the number of paired users. To balance the performance improvement and decoding complexity, two users are selected to implement NOMA, with one of them belonging to group \( \{A_i\} \) and the other to group \( \{B_i\} \). Different from the conventional distance-based and angle-based user pairing strategies (cf. Table I), we propose a user pairing strategy jointly considering the distance and angle information. More specifically, the first user is randomly selected from group \( \{B_i\} \), and it is named as the typical user and denoted as \( U_0 \). The user paired with the typical user \( U_0 \) is selected from group \( \{A_i\} \) denoted as \( U_p \) according to the following minimum angle-difference criterion as

\[
U_p = \arg \min_{U_i \in \Phi_1} |\phi_0 - \phi_i|, \tag{1}
\]

where \( \phi_0 \) and \( \phi_i \) are AOD of \( U_0 \) and \( U_i \) with respect to BS, respectively. \( \phi_0i = |\phi_0 - \phi_i| \) is the absolute difference between the angles \( \phi_0 \) and \( \phi_i \). In a nutshell, the user pairing strategy selects the paired user from group \( \{A_i\} \) who has a minimum angle difference to \( U_0 \). Subsequently, BS switches between single-beam and multi-beam NOMA transmission with flexibility according to the value of angle difference \( \phi_0p \). If the angle difference \( \phi_0p \) is not larger than the beamwidth of main lobe \( \theta_B \) (i.e., \( \phi_0p \leq \theta_B \)), BS performs single-beam NOMA transmission. Otherwise, two analog beams are generated by BS with single RF chain. To elaborate further, Fig. 2 shows the illustration of flexibly switching single-beam and multi-beam NOMA transmission based on the angle difference of paired users in proposed agile-beam NOMA scheme.
C. CHANNEL MODEL

Different from conventional lower frequency cellular networks, the mmWave channel consists of a line-of-sight (LOS) path and several non-line-of-sight (NLOS) paths. Therefore, similar to [26]–[30], the mmWave channel vector can be modeled as

$$
\mathbf{h}_k = \frac{g_k^L a(\phi_k^L)}{\sqrt{1 + d_k^L}} + \sum_{l=1}^{N} \frac{g_k^{NL} l a(\phi_k^{NL})}{\sqrt{1 + d_k^{NL}}}. 
$$

where $g_k^L$ and $\phi_k^L$ represent the small scale fading and the AOD for LOS link, respectively; $g_k^{NL}$ and $\phi_k^{NL}$ represent the small scale fading and the AOD for NLOS link, respectively; $N$ denotes the number of NLOS paths. $\alpha_L$ and $\alpha_{NL}$ are the path loss exponents for LOS path and NLOS path, respectively. In addition, $a(\cdot) \in \mathbb{C}^{M \times 1}$ is a beam steering vector which is defined as

$$
\mathbf{a}(\phi) = \left[ 1, e^{-j\pi \cos(\phi)}, \ldots, e^{-j(M-1)\pi \cos(\phi)} \right]^T.
$$

In mmWave communication systems, a LOS link can achieve 20 dB gain over those NLOS links. Therefore, the LOS path is dominant if it exists, or one of the NLOS paths is dominant if the LOS path does not exist. Considering the characteristics of directionality and spatial sparsity in angle domain in mmWave channel, we adopt the single-path channel model [7], [8] only considering the dominant path from BS to user $k$ expressed as

$$
\mathbf{h}_k = \frac{g_k a(\phi_k)}{\sqrt{1 + d_k^2}}.
$$

where $g_k$ denotes the small scale fading gain. For analytical tractability, we assume that $g_k$ is complex Gaussian distributed, i.e., $g_k \sim \mathcal{CN}(0, 1)$ similarly to [7]–[9]. Moreover, $\sqrt{1 + d_k^2}$ denotes the path loss and $d_k$ is the distance from BS to the user $k$, with $\alpha$ representing the exponent of path loss.

Beamforming with a large antenna array is usually adopted in mmWave NOMA network. The agile-beam NOMA scheme enables BS to serve the two paired users with a single beam or two beams flexibly. In one case, a single beam is generated at BS, and beam vector $\mathbf{w}_s \in \mathbb{C}^{M \times 1}$ of this single beam for user $k$ is defined as

$$
\mathbf{w}_s(\phi_k) = \frac{1}{\sqrt{M}} \left[ 1, e^{-j\pi \cos(\phi_k)}, \ldots, e^{-j(M-1)\pi \cos(\phi_k)} \right]^T,
$$

where $\phi_k$ is AOD from BS to user $k$. Following [31], with the beam pattern mentioned above, the beamwidth $\theta_B$ can be given as $\theta_B = \frac{0.891 \lambda}{M}$ rad, in which $\lambda$ is the wavelength and $d$ denotes the inter-antenna space.

In the other case, a beam splitting technique [12], [13] is utilized, and all antenna elements are divided into two subarrays to generate two analog beams serving user $k$ and user $j$ simultaneously with $M_k$ antenna elements and $M_j$ antenna elements, respectively, satisfying $M_k + M_j = M$. Beam vector $\mathbf{w}_k \in \mathbb{C}^{M_k \times 1}$ for user $k$ is given by

$$
\mathbf{w}_k(\phi_k) = \frac{1}{\sqrt{M}} \left[ 1, e^{-j\pi \cos(\phi_k)}, \ldots, e^{-j(M_k-1)\pi \cos(\phi_k)} \right]^T,
$$

and beam vector for $M_j$ antennas subarray $\mathbf{w}_j \in \mathbb{C}^{M_j \times 1}$ is defined as

$$
\mathbf{w}_j(\phi_j) = \frac{e^{-jM_j\pi \cos(\phi_j)}}{\sqrt{M}} \left[ 1, \ldots, e^{-j(M_j-1)\pi \cos(\phi_j)} \right]^T,
$$

where the term $e^{-jM_j\pi \cos(\phi_j)}$ is used to synchronize the phase of the two subarrays [13]. As a result, the analog precoder for $M$ antennas $\mathbf{w}_m \in \mathbb{C}^{M \times 1}$ is given by

$$
\mathbf{w}_m = \left[ \mathbf{w}_k^T(\phi_k), \mathbf{w}_j^T(\phi_j) \right]^T.
$$

In the 2-user NOMA system, we assume that BS transmits a signal $s_0$ and $s_p$ to $U_0$ and $U_p$, where $\mathbb{E}(|s_0|^2) = 1$ and $\mathbb{E}(|s_p|^2) = 1$, with transmission power allocation coefficients $\beta_0$ and $\beta_p$, respectively. The power allocation coefficients satisfy the conditions that $\beta_0 \geq \beta_p$ and $\beta_0^2 + \beta_p^2 = 1$. According to the superimpose transmission in NOMA system, BS transmits signal $s$ written as

$$
s = \beta_0 s_0 + \beta_p s_p.
$$

And the received signal at $U_0$ and $U_p$ are

$$
y_0 = h_0^T \mathbf{w}_i(\beta_0 s_0 + \beta_p s_p) + n_0,
y_p = h_p^T \mathbf{w}_i(\beta_0 s_0 + \beta_p s_p) + n_p,
$$

where $i \in \{s, m\}$ represents either single-beam ($s$) or multi-beam ($m$) transmission in agile-beam NOMA scheme. $n_0$ and $n_p$ denote the Gaussian white noise with power $\sigma^2$ at $U_0$ and $U_p$, respectively.

III. PERFORMANCE ANALYSIS

BS has an adaptive capability to flexibly switch between the single-beam and multi-beam transmission in the agile-beam NOMA scheme. In this section, we first derive the coverage probability for single-beam and multi-beam transmission cases in detail. Then, aiming to obtain some insights, asymptotic analysis is carried out by using some assumptions, e.g., narrow beamwidth and sufficient antennas. Lastly, the sum rate of the proposed agile-beam NOMA scheme is obtained.
A. COVERAGE PROBABILITY

The agile-beam NOMA scheme enables BS to flexibly serve the two paired users with a single beam or two analog beams. Therefore, the coverage probability are derived separately for these two cases. At first, we derive the coverage probability of paired users in single-beam NOMA transmission.

1) SINGLE-BEAM CASE

When the paired two users are within the main lobe generated by a single beam, BS adjusts its beam orientation so that its boresight direction has the same angle difference with respect to users $U_0$ and $U_p$. Assuming $\phi_0 < \phi_p$, which will be shown no impact on the performance in Section III-B, the boresight of BS can be given by

$$\tilde{\phi} = \frac{\phi_0 + \phi_p}{2} = \phi_0 + \frac{\phi_{0p}}{2}. \quad (11)$$

Beam weighted vector of this single beam can be written as

$$\mathbf{w}_s(\tilde{\phi}) = \frac{1}{\sqrt{M}} \left[ 1, e^{-j2\pi \cos(\tilde{\phi})}, \ldots, e^{-j(M-1)\pi \cos(\tilde{\phi})} \right] ^T. \quad (12)$$

As a result, effective channel gain $|\mathbf{h}_w(\tilde{\phi})|^2$ of user $k$ can be written as

$$|\mathbf{h}_w(\tilde{\phi})|^2 = \frac{\rho_{0k} |\mathbf{a}^H(\phi_0)\mathbf{w}_s(\tilde{\phi})|^2}{1 + d_k^2} F_M(\pi \cos(\phi_k) - \cos(\tilde{\phi})), \quad (13)$$

where $F_M(\cdot)$ represents the Fejér kernel function.

According to decoding order of NOMA, $U_0$ treats its partner’s signal as noise and directly decodes its own information. Therefore, the signal-to-interference-plus-noise ratio (SINR) received at $U_0$ can be expressed as

$$\text{SINR}_0 = \frac{\rho_{0k} |\mathbf{a}^H(\phi_0)\mathbf{w}_s(\tilde{\phi})|^2}{\rho + |\mathbf{a}^H(\phi_0)\mathbf{w}_s(\tilde{\phi})|^2 + 1}, \quad (14)$$

where $\rho$ denotes the ratio of transmit power to noise power, i.e., the signal-to-noise ratio (SNR). Moreover, the coverage probability for $U_0$ means probability of decoding its own message successfully and it can be written as

$$P_{0,s} = \mathbb{P}(\log_2(1 + \text{SINR}_0) \geq R_0) = \mathbb{P}(\text{SINR}_0 \geq \epsilon_0), \quad (15)$$

where $R_0$ and $\epsilon_0$ are rate threshold and SINR threshold ($\epsilon_0 = 2^{R_0} - 1$) for $U_0$, respectively.

In the other hand, $U_p$ implements SIC procedure and first tries to decode its partner’s information with

$$\text{SINR}_{p \rightarrow 0} = \frac{\rho_{0p} |\mathbf{a}^H(\phi_0 + \phi_{0p})\mathbf{w}_s(\tilde{\phi})|^2}{\rho + |\mathbf{a}^H(\phi_0 + \phi_{0p})\mathbf{w}_s(\tilde{\phi})|^2 + 1}. \quad (16)$$

If $\text{SINR}_{p \rightarrow 0} \geq \epsilon_0$, which means that $U_p$ decodes its partner’s message successfully, and then $U_p$ can remove the information of $U_0$ and decode its own message. With the aid of SIC, the SINR received at $U_p$ can be written as

$$\text{SINR}_p = \rho \frac{\rho_{0p} |\mathbf{a}^H(\phi_0 + \phi_{0p})\mathbf{w}_s(\tilde{\phi})|^2}{1 + d_k^2} \cdot \rho_{0p}^2 \rho_p, \quad (17)$$

Therefore, coverage event for $U_p$ occurs when $U_p$ decodes its partner’s and its own message successfully. In other words, coverage probability experienced by $U_p$ can be expressed as

$$P_{p,s} = \mathbb{P}(\text{SINR}_{p \rightarrow 0} \geq \epsilon_0, \text{SINR}_p \geq \epsilon_p), \quad (18)$$

where $R_p$ is rate threshold required by $U_p$ and $\epsilon_p = 2^{R_p} - 1$.

Due to the randomness of the small scale fading, the angle of $U_0$, the angle difference between the $U_0$ and $U_p$, as well as the distances from BS to users, the SINRs given in (14), (16) and (17) are all random variables. In order to calculate coverage probability, the first thing is to derive probability density functions (PDFs) of these variables, including the user distance and user angle. The corresponding results are given in the following lemmas.

Lemma 1: The PDF of distance from BS to $U_p$ and $U_0$ are respectively given by

$$f_{d} (r) = \frac{2r}{R_1^2}, \quad 0 \leq r \leq R_1, \quad (19)$$

$$f_{d} (r) = \frac{2r}{R_2^2 - R_1^2}, \quad R_2 \leq r \leq R_3, \quad (20)$$

Proof. Please refer to Appendix A.

The spatial locations of users in group $\{A_i\}$ form homogeneous PPP $\Phi_1$ with density $\lambda_1$. Therefore, the number of users in group $\{A_i\}$ obeys Poisson distribution and its expression can be written as

$$\mathbb{P}(K = k) = \left( \frac{\lambda_1 R_1^2}{k!} \right) e^{-\lambda_1 R_1^2}, \quad (21)$$

We assume that there is at least one user in near user group to ensure successful pairing. Hence, $k \geq 1, k \in \mathbb{Z}$. In addition, the AOD of user selected randomly from group $\{A_i\}$ or $\{B_j\}$ is uniformly distributed over the interval $[0, 2\pi]$, and then PDF of the random variable $\phi_0$ can be expressed as $f_{\phi_0}(\phi) = \frac{1}{\pi}$, $0 \leq \phi \leq 2\pi$. Based on the PDF of $\phi_0$ and equation (21), we can obtain the PDF of $\phi_{0p}$ as follows.

Lemma 2: According to the user pairing strategy, the PDF of angle difference $\phi_{0p}$ between the typical user $U_0$ and the paired user $U_p$ can be derived as

$$f_{\phi_{0p}}(\phi) = \frac{K}{\pi} \left( \frac{2\pi - \varphi}{2\pi} \right)^{2K-1}, \quad \varphi \in [0, 2\pi], \quad (22)$$

which is conditioned on that the number of users in group $\{A_i\}$ is $K$.

Proof. Please refer to Appendix B.

Proposition 1: When angle difference $\phi_{0p}$ is not larger than beamwidth $\theta_b$, BS switches to single-beam transmission of agile-beam NOMA scheme. Based on the PDFs of distance and angle difference in (19), (20) and (22), the coverage probability of $U_0$ and $U_p$ can be derived as

$$P_{0,s} = \int_{R_2}^{R_1} \int_{0}^{2\pi} \int_{0}^{\theta_b} \rho \frac{\rho_{0p} |\mathbf{a}^H(\phi_0 + \phi_{0p})\mathbf{w}_s(\tilde{\phi})|^2}{1 + d_k^2} \cdot \rho_{0p}^2 \rho_p \times e^{-\frac{F_M(\pi \cos(\phi) - \cos(\tilde{\phi}))}{\rho}} f_{\phi_{0p}}(\phi) f_{\phi_0}(\phi) f_0(r) \, d\phi \, dr \, dr, \quad (23)$$
\[ P_{p,s} = \int_{R_1} \int_{0}^{2\pi} \int_{0}^{\theta_B} e^{\frac{f_{\phi_0}(\phi)f_{\phi}(\phi)p(r)}{e^{\eta_1(1+\nu)}}} \, d\phi \, d\rho \, dr. \]

if \( \beta_0^2 - \epsilon_0 \beta_p^2 > 0 \); Otherwise \( P_{p,s} = 0 \) and \( P_{p,0} = 0 \), where \( \eta_0 = \frac{\epsilon_0}{\rho \beta_0^2 - \epsilon_0 \beta_p^2} \), \( \eta_0 = \max \left\{ \frac{\epsilon_0}{\rho \beta_0^2 - \epsilon_0 \beta_p^2}, \frac{\epsilon_p}{\rho \beta_p^2} \right\} \)

Proof. Please refer to Appendix C.

\[ P_{0,m} = \int_{R_2} \int_{0}^{2\pi} \int_{0}^{\theta_B} e^{\frac{f_{\phi_0}(\phi)f_{\phi}(\phi)p(r)}{e^{\eta_1(1+\nu)}}} \, d\phi \, d\rho \, dr \]

conditioned on \( \beta_0^2 - \epsilon_0 \beta_p^2 > 0 \); otherwise \( P_{0,m} = 0 \) and \( P_{p,m} = 0 \).

2) MULTI-BEAM CASE
In the agile-beam NOMA scheme, when the paired users cannot be served by single-beam NOMA transmission, BS adaptively switches to multi-beam transmission. More specifically, BS adopts the beam splitting technique to generate two analog beams with a single RF chain. As mentioned above, the beam weighted vector of two beams for \( U_0 \) and \( U_p \) generated by a RF chain can be written as

\[ \mathbf{w}_m = \left\{ \mathbf{w}_0(\phi_0), \mathbf{w}_p(\phi_p) \right\}^T, \]

where \( \mathbf{w}_m \in \mathbb{C}^{M \times 1}, \mathbf{w}_0(\phi_0) \in \mathbb{C}^{M_0 \times 1}, \mathbf{w}_p(\phi_p) \in \mathbb{C}^{M_p \times 1} \), \( \mathbf{w}_0(\phi_0) = \frac{1}{\sqrt{M_0}} \left[ 1, \ldots, e^{-jM_0 - 1} \cos(\phi_0) \right]^T \) and \( \mathbf{w}_p(\phi_p) = e^{-jM_p \cos(\phi_p)} \left[ 1, \ldots, e^{-jM_p - 1} \cos(\phi_p) \right]^T \). \( M_0 \) and \( M_p \) denote the number of allocated antennas for \( U_0 \) and \( U_p \), respectively, satisfying \( M_0 + M_p = M \).

Similar to single-beam NOMA transmission, BS employs superimpose transmission, and then the typical user \( U_0 \) decodes its own information while the paired user \( U_p \) implements SIC procedure. The corresponding SINR expressions are given as follows.

The SINR received at \( U_0 \) can be written as

\[ \text{SINR}_0 = \frac{\|g_0\|^2}{\eta_0} \left| \mathbf{a}_H(\phi_0) \mathbf{w}_0 \right|^2 \beta_0^2, \]

and \( U_p \) decodes the information of \( U_0 \) with

\[ \text{SINR}_{p \rightarrow 0} = \frac{\|g_0\|^2}{\eta_0} \left| \mathbf{a}_H(\phi_p) \mathbf{w}_0 \right|^2 \beta_0^2, \]

After SIC procedure, the SINR received at \( U_p \) can be given by

\[ \text{SINR}_p = \rho \frac{\|g_p\|^2}{1 + d_p} \left| \mathbf{a}_H(\phi_p) \mathbf{w}_m \right|^2 \beta_p^2. \]

Through same derivation as single-beam case, the coverage probability of \( U_0 \) and \( U_p \) in multi-beam NOMA transmission are given in the following proposition.

**Proposition 2:** When angle difference \( \phi_0 \), of paired users is larger than \( \theta_B \), the agile-beam NOMA scheme enables BS to switch to multi-beam transmission adaptively. Based on the

PDFs in (19), (20) and (22), the coverage probability of \( U_0 \) and \( U_p \) can be derived as

\[ P_{0,m} = \int_{R_2} \int_{0}^{2\pi} \int_{0}^{\theta_B} e^{\frac{f_{\phi_0}(\phi)f_{\phi}(\phi)p(r)}{e^{\eta_1(1+\nu)}}} \, d\phi \, d\rho \, dr \]

conditioned on \( \beta_0^2 - \epsilon_0 \beta_p^2 > 0 \); otherwise \( P_{0,m} = 0 \) and \( P_{p,m} = 0 \).

Proof. Following similar steps of deriving the coverage probability at \( U_0 \) and \( U_p \) in single-beam NOMA transmission case, we can obtain the coverage probability of \( U_0 \) and \( U_p \) in multi-beam transmission case. Therefore, we omit the detailed proof.

**B. ASYMPTOTIC ANALYSIS**
The obtained coverage probability in (23), (24), (29) and (30) are quite complicated, which involve the calculation of triple integrals. In order to obtain some insights, some approximation and asymptotic analysis is carried out in single-beam and multi-beam cases.

1) SINGLE-BEAM CASE
Based on user pairing result \( \phi_0 \leq \theta_B \), BS switches to single-beam NOMA transmission and generates a single beam to communicate with the paired users. The obtained coverage probability expressed in (23) and (24) incorporates the Fejér kernel function, and it is very difficult to calculate the triple integrals involved in these expressions. Aiming to obtain some insights, some approximation and asymptotic analysis are conducted by using some assumptions, e.g., narrow beamwidth and sufficient antennas.

Following the idea of [7], the beamforming gain at \( U_0 \) in single-beam NOMA case can be approximated as

\[ \left| \mathbf{a}_H(\phi_0) \mathbf{w}(\phi) \right|^2 \]

\[ = F_M \left( \pi \left( \cos(\phi_0) - \cos \left( \frac{\phi_0 + \phi_p}{2} \right) \right) \right) \]

\[ \approx M \left( 1 - \frac{\pi^2 M^2 \left( \cos \left( \frac{\phi_0 + \phi_p}{2} - \cos(\phi_0) \right) \right)}{12} \right). \]

where the approximation follows from that beam generated by BS with a single RF chain is very narrow. In addition, from (31), it can be observed that the assumption of \( \phi_0 < \phi_p \) has no impact on the beamforming gain. The approximation of beamforming gain at \( U_p \) follows similar steps. Then substituting the approximation expression of beamforming gain into (23) and (24), the approximate expressions of coverage probability can be calculated.
Proposition 3: Based on the approximation of beamforming gain and after some algebraic manipulations when path loss exponent $\alpha = 2$, the coverage probability of $U_0$ and $U_p$ in single-beam NOMA transmission case can be derived as

$$P_{0,s} \approx c_1 - \frac{\eta_0}{M} \left(1 + \frac{R_3^2 + R_2^2}{2} \right) \left(c_1 + \frac{\pi^2 M^2}{96} c_2\right),$$  \hspace{1cm} (32)

$$P_{p,s} \approx c_1 - \frac{\eta_p}{M} \left(1 + \frac{R_1^2}{2} \right) \left(c_1 + \frac{\pi^2 M^2}{96} c_2\right),$$  \hspace{1cm} (33)

if $\beta_0^2 - \epsilon_0 \beta_p^2 > 0$; otherwise $P_{0,s} = 0$ and $P_{p,s} = 0$, where $c_1 = \sum_k \mathbb{P}(K = k) f_0^h f_{\phi_0}(\varphi)d\varphi$, and $c_2 = \sum_k \mathbb{P}(K = k) \int_0^\varphi \varphi^2 f_{\phi_0}(\varphi)d\varphi$.

Proof. Please refer to Appendix D. \hfill \blacksquare

2) MULTI-BEAM CASE

When BS switches to multi-beam NOMA transmission, it is very difficult to express the beamforming gain $|\mathbf{a}^H \mathbf{w}_m|^2$ with a simple analytical expression and calculate the integral in (29), (30). Alternatively, we follow the idea of [12] and adopt asymptotic performance analysis. We focus on the case when the antennas equipped at BS and the number of antennas allocated to $U_0$ and $U_p$ are sufficiently large. As a result, the beamforming gain $|\mathbf{a}^H \mathbf{w}_m|^2$ can be approximated by

$$|\mathbf{a}^H (\phi_0) \mathbf{w}_m|^2 \approx \frac{M_0^2}{M},$$

$$|\mathbf{a}^H (\phi_p) \mathbf{w}_m|^2 \approx \frac{M_p^2}{M}. \hspace{1cm} (34)$$

Proposition 4: Based on the approximation of beamforming gain in multi-beam NOMA transmission and path loss exponent $\alpha = 2$, the coverage probability of $U_0$ and $U_p$ in the case of multi-beam NOMA can be derived as

$$P_{0,m} \approx \frac{1 - c_1 e^{-\frac{\eta_0 M_0^2}{M_0}}}{(R_3^2 + R_2^2)} \frac{e^{-\frac{\eta_0 M_0^2}{M_0}} - e^{-\frac{M_0 R_2^2}{M_0}}}{R_2^2 \eta_0 \frac{M_0}{M}} \left(c_1 + \frac{\pi^2 M^2}{96} c_2\right). \hspace{1cm} (35)$$

$$P_{p,m} \approx \frac{1 - c_1 e^{-\frac{\eta_p M_p^2}{M_p}}}{R_1^2 \eta_p \frac{M_p}{M}} \left(1 - e^{-\frac{M_p R_1^2}{M_p}}\right) \left(c_1 + \frac{\pi^2 M^2}{96} c_2\right). \hspace{1cm} (36)$$

on condition of $\beta_0^2 - \epsilon_0 \beta_p^2 > 0$; otherwise $P_{0,m} = 0$ and $P_{p,m} = 0$.

Proof: Based on the approximation of beamforming gain in multi-beam NOMA transmission, substituting the PDFs of distance, angle $\phi_0$, angle difference $\phi_p$ and path loss exponent $\alpha = 2$ into (29), (30), after some algebraic manipulations, we can obtain the coverage probability of $U_0$ and $U_p$ in (35) and (36), respectively, and thus we omit the detailed proof. \hfill \blacksquare

C. SUM RATE

Theorem 1: The coverage probability of $U_0$ and $U_p$ of the proposed agile-beam NOMA scheme in mmWave networks can be expressed as

$$P_0 = P_{0,s} + P_{0,m}, \hspace{1cm} (37)$$

$$P_p = P_{p,s} + P_{p,m}. \hspace{1cm} (38)$$

The sum rate of the paired users is given by

$$R_{\text{sum}} = P_0 R_0 + P_p R_p. \hspace{1cm} (39)$$

Proof: By considering the two cases of single-beam and multi-beam NOMA transmission, we directly obtain the coverage probability of the paired NOMA users and the sum rate of agile-beam NOMA system. \hfill \blacksquare

IV. A MULTIPLE-RF CHAIN SYSTEM

In previous sections, we focus on investigating the agile-beam NOMA scheme in a single single-RF chain system where BS serves a pair of users with a single RF chain. In this section, we extend to multiple RF chains system, where BS communicates with multiple pairs of users with multiple RF chain and each RF chain serves two users in agile-beam NOMA scheme, as shown in Fig. 3. Consider a scenario where BS equipped with $N_{RF}$ RF chains accessing all $M$ antennas communicates with $K$ users selected by the previous pairing strategy. We assume that BS is in full load, i.e., $K = 2N_{RF}$. We focus on a typical pair of users, referred to as $U_{1,0}$ and $U_{1,p}$ served by RF chain 1, and we assume $U_{1,p}$ is the strong user and $U_{1,0}$ is the weak user.

The received signal at $U_{1,0}$ is given by

$$y_{1,0} = h_{1,0}^H \mathbf{w}_1 \beta_{1,0} s_{1,0} + h_{1,0}^H \mathbf{w}_1 \beta_{1,p} s_{1,p} + \sum_{i=2}^{N_{RF}} h_{1,0}^H \mathbf{w}_i (\beta_{i,0}s_{i,0} + \beta_{i,p}s_{i,p}) + n_{1,0}, \hspace{1cm} (40)$$

where the first term represents the desired signal of $U_{1,0}$, the second term denotes the intra-group interference caused by the paired user $U_{1,p}$, and the third term represents the intergroup interference caused by all other RF chains. $\mathbf{w}_i \in \mathbb{C}^{M \times 1}$ can be the beam vector in single-beam case or in multi-beam case according to the phase difference of paired users $U_{1,0}$ and $U_{1,p}$. Therefore, $U_{1,0}$ directly decodes its own signal with the following SINR:

$$\text{SINR}_{1,0} = \frac{|h_{1,0}^H \mathbf{w}_1|^2 \beta_{1,0}^2}{|h_{1,0}^H \mathbf{w}_1|^2 \beta_{1,p}^2 + \sum_{i=2}^{N_{RF}} |h_{1,0}^H \mathbf{w}_i|^2 + 1/p}. \hspace{1cm} (41)$$
The SINR for $U_{1,p}$ to decode its partner’s information can be expressed as

$$SINR_{1,p} = \frac{|h^H_{1,p}w|_2^2 \beta^2_{1,p} \sum_{i=2}^{NRF} |h^H_{i,p}w|_2^2 + \frac{1}{\rho}}{\sum_{i=2}^{NRF} |h^H_{i,p}w|_2^2 + \frac{1}{\rho}}.$$ (42)

After performing SIC, the SINR for $U_{1,p}$ to decode its own message can be expressed as

$$SINR_{1,p} = \frac{|h^H_{1,p}w|_2^2 \beta^2_{1,p}}{\sum_{i=2}^{NRF} |h^H_{i,p}w|_2^2 + \frac{1}{\rho}}.$$ (43)

Therefore, based on the definition of coverage probability, the coverage probability for $U_{1,0}$ can be written as

$$P_{1,0} = P(SINR_{1,0} \geq \epsilon_{1,0}).$$ (44)

Similarly, the coverage probability for $U_{1,p}$ can be given as

$$P_{1,p} = P(SINR_{1,p} \to 0 \geq \epsilon_{1,0}, SINR_{1,p} \geq \epsilon_{1,p}).$$ (45)

The sum rate of this typical user pair can be expressed as

$$R_{\text{sum}} = P_{1,0}R_{1,0} + P_{1,p}R_{1,p},$$ (46)

where $R_{1,0}$ and $R_{1,p}$ are the rate thresholds for $U_{1,0}$ and $U_{1,p}$, respectively, and $\epsilon_{1,0} = 2R_{1,0} - 1, \epsilon_{1,p} = 2R_{1,p} - 1$.

Because of the existence of inter-group interference, the coverage probability and sum rate of this typical user pair are difficult to calculate with the tool of stochastic geometry. Fortunately, when the analog beamforming structure is employed at BS, the effective channel gain between $U_{1,k}$ ($k \in \{0, p\}$) and all the other RF chains, $\forall i \neq 1, |h^H_{1,k}w_i|_2^2$, is generally small, and the inter-group interference is limited. When BS employs hybrid beamforming structure, ZF precoding can be utilized to manage the inter-group interference. We provide the Monte Carlo simulation results of coverage probability and sum rate in the multiple-RF chain system to future work.

**V. PERFORMANCE EVALUATION**

In this section, we evaluate the performance of agile-beam NOMA scheme and compare it with conventional NOMA and OMA schemes. For conventional NOMA, if angle difference paired users is not larger than beamwidth of single beam, BS adopts single-beam NOMA transmission, otherwise, employing OMA transmission. In OMA, no matter how close the paired users deploy, typical user and the paired user are served in the first and second halves of a time slot, respectively. In simulation, the number of antennas equipped at BS is $M = 128$. Following the system parameter setting in [7], [8], the radius of disk is set as $R_1 = 2.5$ m, and distributed density of this disk is $\lambda_1 = 2$. The inner and outer radius of ring are set as $R_2 = 8$ m and $R_3 = 10$ m, respectively, with distributed density of $\lambda_2 = 1$.

Fig. 4 and Fig. 5 show the impact of SNR on coverage probability and sum rate in agile-beam NOMA, conventional NOMA and OMA schemes. Monte Carlo simulation results (sim), analytical results (ana) and approximate results (app) are provided. In Fig. 4(a) and Fig. 4(b), with the increase of SNR, coverage probability of two paired users increase, and the gap between approximate results and analytical results decreases. Monte Carlo simulation results anastomose with
analytical analysis, which validates the performance analysis in this paper. As can be observed from Fig. 4.(a), with different antenna allocation, coverage probability of typical user $U_0$ presents different performance. Specifically, when the number of antennas allocated to $U_p$ is $M_p = 100$, coverage probability of $U_0$ in agile-beam NOMA scheme is lower than coverage probability in conventional NOMA and OMA scheme. However, when increasing antennas allocated to $U_0$, i.e., $M_p = 50$, coverage probability of $U_0$ is superior to the other two multiple access schemes because of the increase of beamforming gain. From Fig. 4.(b), we can observe that coverage probability of $U_p$ is superior to other two schemes under $M_p = 50$ and $M_p = 100$. When the number of antenna allocated to $U_p$ decreases, coverage probability of $U_p$ in agile-beam NOMA declines but superior to other schemes. The decline is resulted from the decrease of beamforming gain, and the reason why coverage probability is still superior to other schemes is the implementation of SIC during multi-beam transmission in agile-beam NOMA scheme, which cancels the interference from $U_0$. Moreover, it is worth pointing out that appropriate antenna allocation is significant to ensure coverage of two paired users in our proposed agile-beam NOMA scheme.

In Fig. 5, the label ‘MPC’ represents the sum rate with MPC channel in agile-beam NOMA scheme. As can be observed from Fig. 5, sum rate increases with the increase of SNR. The proposed agile-beam NOMA scheme can yield a significant sum rate gain over conventional NOMA scheme and OMA scheme, which validates the superiority of agile-beam NOMA scheme over the other two MA scheme. In addition, the sum rate obtained from asymptotic analysis are tightly matching with numerical simulation results at high SNR. Moreover, the sum rate with MPC channel anastomoses with that of single-path channel model, which justifies the assumption of single-path channel model in this paper.

A sophisticated user pairing strategy, RNRF [22], is adopted and compared with the proposed user pairing strategy in terms of coverage probability. The RNRF strategy selects one user from group $A_i$ randomly as the strong user and pairs it with another user selected from group $B_i$ randomly as the weak user. For the sake of rigor, the paired users selected by the RNRF pairing strategy are served by BS with the proposed agile-beam NOMA scheme. The comparison between the proposed user pairing strategy and the RNRF strategy is presented in Fig. 6. The label ‘RNRF’ represents the coverage probability of users with RNRF strategy. It can be observed that the coverage probability of $U_0$ and $U_p$ are larger than that of RNRF strategy and thus the sum rate of the proposed user pairing strategy is superior to RNRF strategy. The proposed user pairing strategy utilizes the angle information to facilitate the switching between single-beam transmission and multi-beam transmission, which improves the coverage probability and sum rate significantly.

As mentioned above, antenna resource allocation is vital to coverage probability and sum rate in agile-beam NOMA scheme. We investigate the impact of antennas allocated to $U_p$ on coverage probability with different power allocation coefficients in Fig. 7. As can be observed from Fig. 7, when power coefficient of $U_0$ is $\beta_0 = 0.9$, with the increase of the antennas allocated to $U_p$, coverage probability of $U_0$ decreases even lower than the OMA scheme. When $\beta_0 = 0.75$, coverage probability of $U_0$ is always inferior to OMA scheme no matter how many antennas are allocated to $U_p$. Moreover, with the increase of antennas allocated to $U_p$, i.e., the increase of beamforming gain at $U_p$, coverage probability of $U_p$ raises and is superior to OMA. The more antennas are allocated to user, the closer approximate results are to analytical results, which is the reason why approximate coverage probability of $U_0$ deviates from analytical results with the increase of $M_p$ while approximate coverage probability of $U_p$ becomes closer to analytical results.

The comparison of coverage probability between the single-RF chain system and the multi-RF chain system is presented in Fig. 8. The number of RF chains in multi-RF chain system is set as $N_{RF} = 4$. In Fig. 9, we provide the
comparison of sum rate between the single-RF chain system and the multi-RF chain system. The sum rates of 2-RF chain system and 4-RF chain system are provided to investigate the impact of the number of RF chains on the sum rate. From Fig. 8 and Fig. 9, we can observe that the multi-RF chain system presents the same trend as the single-RF chain system with the increase of SNR in terms of coverage probability and sum rate. The coverage probability and sum rate of multi-RF chain system are inferior to the ones of single-RF chain system and 4-RF chain system are provided to investigate the comparison of sum rate between the single-RF chain system and the multi-RF chain system. The sum rates of 2-RF chain system and 4-RF chain system are provided to investigate the impact of the number of RF chains on the sum rate. From Fig. 8 and Fig. 9, we can observe that the multi-RF chain system presents the same trend as the single-RF chain system with the increase of SNR in terms of coverage probability and sum rate. The coverage probability and sum rate of multi-RF chain system are inferior to the ones of single-RF chain system and 4-RF chain system are provided to investigate the comparison of sum rate between the single-RF chain system and the multi-RF chain system. The sum rates of 2-RF chain system and 4-RF chain system are provided to investigate the impact of the number of RF chains on the sum rate. From Fig. 8 and Fig. 9, we can observe that the multi-RF chain system presents the same trend as the single-RF chain system with the increase of SNR in terms of coverage probability and sum rate. The coverage probability and sum rate of multi-RF chain system are inferior to the ones of single-RF chain system and 4-RF chain system are provided to investigate the comparison of sum rate between the single-RF chain system and the multi-RF chain system. The sum rates of 2-RF chain system and 4-RF chain system are provided to investigate the impact of the number of RF chains on the sum rate. From Fig. 8 and Fig. 9, we can observe that the multi-RF chain system presents the same trend as the single-RF chain system with the increase of SNR in terms of coverage probability and sum rate. The coverage probability and sum rate of multi-RF chain system are inferior to the ones of single-RF chain system and 4-RF chain system are provided to investigate the comparison of sum rate between the single-RF chain system and the multi-RF chain system. The sum rates of 2-RF chain system and 4-RF chain system are provided to investigate the impact of the number of RF chains on the sum rate. From Fig. 8 and Fig. 9, we can observe that the multi-RF chain system presents the same trend as the single-RF chain system with the increase of SNR in terms of coverage probability and sum rate. The coverage probability and sum rate of multi-RF chain system are inferior to the ones of single-RF chain system and 4-RF chain system are provided to investigate the comparison of sum rate between the single-RF chain system and the multi-RF chain system. The sum rates of 2-RF chain system and 4-RF chain system are provided to investigate the impact of the number of RF chains on the sum rate. From Fig. 8 and Fig. 9, we can observe that the multi-RF chain system presents the same trend as the single-RF chain system with the increase of SNR in terms of coverage probability and sum rate. The coverage probability and sum rate of multi-RF chain system are inferior to the ones of single-RF chain system and 4-RF chain system are provided to investigate the comparison of sum rate between the single-RF chain system and the multi-RF chain system. The sum rates of 2-RF chain system and 4-RF chain system are provided to investigate the impact of the number of RF chains on the sum rate. From Fig. 8 and Fig. 9, we can observe that the multi-RF chain system presents the same trend as the single-RF chain system with the increase of SNR in terms of coverage probability and sum rate. The coverage probability and sum rate of multi-RF chain system are inferior to the ones of single-RF chain system and 4-RF chain system are provided to investigate the comparison of sum rate between the single-RF chain system and the multi-RF chain system. The sum rates of 2-RF chain system and 4-RF chain system are provided to investigate the impact of the number of RF chains on the sum rate. From Fig. 8 and Fig. 9, we can observe that the multi-RF chain system presents the same trend as the single-RF chain system with the increase of SNR in terms of coverage probability and sum rate. The coverage probability and sum rate of multi-RF chain system are inferior to the ones of single-RF chain system and 4-RF chain system are provided to investigate the comparison of sum rate between the single-RF chain system and the multi-RF chain system. The sum rates of 2-RF chain system and 4-RF chain system are provided to investigate the impact of the number of RF chains on the sum rate. From Fig. 8 and Fig. 9, we can observe that the multi-RF chain system presents the same trend as the single-RF chain system with the increase of SNR in terms of coverage probability and sum rate. The coverage probability and sum rate of multi-RF chain system are inferior to the ones of single-RF chain system and 4-RF chain system are provided to investigate the comparison of sum rate between the single-RF chain system and the multi-RF chain system. The sum rates of 2-RF chain system and 4-RF chain system are provided to investigate the impact of the number of RF chains on the sum rate. From Fig. 8 and Fig. 9, we can observe that the multi-RF chain system presents the same trend as the single-RF chain system with the increase of SNR in terms of coverage probability and sum rate. The coverage probability and sum rate of multi-RF chain system are inferior to the ones of single-RF chain system and 4-RF chain system are provided to investigate the comparison of sum rate between the single-RF chain system and the multi-RF chain system. The sum rates of 2-RF chain system and 4-RF chain system are provided to investigate the impact of the number of RF chains on the sum rate. From Fig. 8 and Fig. 9, we can observe that the multi-RF chain system presents the same trend as the single-RF chain system with the increase of SNR in terms of coverage probability and sum rate. The coverage probability and sum rate of multi-RF chain system are inferior to the ones of single-RF chain system and 4-RF chain system are provided to investigate the comparison of sum rate between the single-RF chain system and the multi-RF chain system. The sum rates of 2-RF chain system and 4-RF chain system are provided to investigate the impact of the number of RF chains on the sum rate. From Fig. 8 and Fig. 9, we can observe that the multi-RF chain system presents the same trend as the single-RF chain system with the increase of SNR in terms of coverage probability and sum rate. The coverage probability and sum rate of multi-RF chain system are inferior to the ones of single-RF chain system and 4-RF chain system are provided to investigate the comparison of sum rate between the single-RF chain system and the multi-RF chain system.
After taking the first derivative, the PDF of the angle difference \( \phi_{op} \) can be given in (22). The lemma 2 is proved.

**APPENDIX C**

**PROOF FOR PROPOSITION 1**

Based on the PDFs of the distance and angle difference \( \phi_{op} \), the coverage probability of \( U_0 \) can be derived as

\[
P_{0,s} = \mathbb{P}(\text{SINR}_0 \geq \epsilon_0) = \mathbb{P}\left( |g|^{2} \geq \frac{\eta_0(1+d_0^2)}{F_M(\pi \cos(\phi_{op}) - \cos(\phi + \frac{1}{2} \phi_{op}))}\right),
\]

where \( \eta_0 = \frac{\epsilon_0}{\rho(\beta^2_0 - \epsilon_0 \beta^2_p)} \) and \( \beta^2_0 - \epsilon_0 \beta^2_p > 0 \). \( \beta^2_0 - \epsilon_0 \beta^2_p \leq 0 \) indicates that the typical user \( U_0 \) cannot decode its own information successfully, and \( P_{0,s} = 0 \). Using the definition of exponential distribution and plugging the PDFs of the distance and angle difference in (20) and (22) into (52), the coverage probability at \( U_0 \) can be obtained in (23). Following similar steps, the coverage probability of \( U_p \) can be derived as

\[
P_{p,s} = \mathbb{P}(\text{SINR}_{0-p} \geq \epsilon_0, \text{SINR}_p \geq \epsilon_p) = \mathbb{P}\left( |g_p|^{2} \geq \frac{\eta_p(1+d_p^2)}{F_M(\pi \cos(\phi_{op} + \phi_{op}) - \cos(\phi + \frac{1}{2} \phi_{op}))}\right).
\]

where \( \eta_p = \max\left\{ \frac{\epsilon_p}{\rho(\beta^2_0 - \epsilon_0 \beta^2_p)}, \frac{\epsilon_p}{\rho \beta^2_p} \right\} \) and \( \beta^2_0 - \epsilon_0 \beta^2_p > 0 \). Similar to the typical user, the paired user \( U_p \) can not decode its partner’s information successfully when \( \beta^2_0 - \epsilon_0 \beta^2_p \leq 0 \). Substituting the PDFs in (19) and (22) into (53), we obtain the coverage probability of \( U_p \) in (24). The proposition 1 is proved.

**APPENDIX D**

**PROOF FOR PROPOSITION 3**

When the number of antennas equipped at BS is sufficiently large, the beamwidth is small, i.e., \( \theta_d \to 0 \), which means \( \phi_{op} \to 0 \) when single-beam NOMA transmission. Based on the triangular function sum and difference product formula, we can obtain

\[
\cos(\phi_0 + \frac{\phi_{op}}{2}) - \cos(\phi_0) \\
\approx -2 \sin(\phi_0 + \frac{\phi_{op}}{4}) \sin(\frac{\phi_{op}}{4}) \\
\approx -\frac{1}{2} \phi_{op} \sin(\phi_0).
\]

The first approximation follows from sum and difference product formula of triangular function and the second approximation follows from \( \sin(x) \approx x \) for \( x \to 0 \). Substituting the approximation in (54) into the beamforming gain expressed in (31), we can obtain

\[
|\mathbf{a}^H(\phi_0)\mathbf{w}(\phi)|^2 \approx M \left( 1 - \frac{\pi^2 M^2 \phi_{op}^2 \sin(\phi_0)^2}{48} \right).
\]

Therefore, the coverage probability at \( U_0 \) can be approximated as follows:

\[
P_{0,s} \approx \int_{R_2}^{R_1} \int_{0}^{2\pi} \int_{0}^{\pi} e^{- \frac{\eta_0(1+d_0^2)}{M(1 - \pi^2 M^2 \phi_{op}^2 \sin(\phi_0)^2)}} d\phi \, d\phi \, d\theta \\
\approx \int_{R_2}^{R_1} \int_{0}^{2\pi} \int_{0}^{\pi} e^{- \frac{\eta_0(1+d_0^2)}{M(1 + \pi^2 M^2 \phi_{op}^2 \sin(\phi_0)^2)}} d\phi \, d\phi \, d\theta \\
\approx \int_{R_2}^{R_1} \int_{0}^{2\pi} \int_{0}^{\pi} \left( 1 - \frac{\pi^2 M^2 \phi_{op}^2 \sin(\phi_0)^2}{48} \right) \, d\phi \, d\phi \, d\theta. \quad (56)
\]

where \( \eta_0 = \frac{\epsilon_0}{\rho(\beta^2_0 - \epsilon_0 \beta^2_p)} \), the second approximation follows from \( \frac{1}{1+x} \approx 1 + x \) for \( x \to 0 \) and the last approximation follows from \( e^{-x} \approx 1 - x \) for \( x \to 0 \) at high SNR, i.e., \( \eta_0 \to 0 \). By substituting the PDFs of distance \( r \), angle of \( U_0 \), angle difference \( \phi_{op} \) and \( \alpha = 2 \) into (56), after some algebraic manipulations, the coverage probability of \( U_0 \) can be expressed in (32). Following similar steps, the coverage probability of \( U_p \) can be obtained in (33). The proof is complete.

**REFERENCES**

[1] K. Chandra, A. S. Marcano, S. M. Mumtaz, R. V. Prasad, and H. L. Christiansen, “Unveiling capacity gains in ultradense networks: Using mm-Wave NOMA,” IEEE Veh. Technol. Mag., vol. 13, no. 2, pp. 75–83, Jun. 2018.

[2] Z. Xiao, L. Dai, Z. Ding, J. Choi, P. Xia, and X. Xia, “Millimeter-wave communication with non-orthogonal multiple access for 5G,” 2017, arXiv:1709.07980. [Online]. Available: https://arxiv.org/abs/1709.07980

[3] S. M. R. Islam, N. Avazov, O. A. Dobre, and K. Kwak, “Power-domain non-orthogonal multiple access (NOMA) in 5G system: Potentials and challenges,” IEEE Commun. Surveys Tuts., vol. 19, no. 2, pp. 721–742, Sep. 2017.

[4] L. Dai, B. Wang, Y. Yuan, S. Han, C.-L. I, and Z. Wang, “Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends,” IEEE Commun. Mag., vol. 53, no. 9, pp. 74–81, Sep. 2015.

[5] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. Bhargava, “A survey on non-orthogonal multiple access for 5G networks: Research challenges and future trends,” IEEE J. Sel. Areas Commun., vol. 35, no. 10, pp. 2181–2195, Oct. 2017.

[6] A. J. Morgado, K. M. S. Huq, J. Rodriguez, C. Politis, and H. Gacanie, “Hybrid resource allocation for millimeter-wave NOMA,” IEEE Wireless Commun., vol. 24, no. 5, pp. 23–29, Oct. 2017.

[7] Z. Ding, P. Fan, and H. V. Poor, “Random beamforming in millimeter-wave NOMA networks,” IEEE Access, vol. 5, pp. 7667–7681, 2017.

[8] T. Lv, Y. Ma, J. Zeng, and P. T. Mathiopoulos, “Millimeter-wave noma transmission in cellular M2M communications for Internet of Things,” IEEE Internet Things J., vol. 5, no. 3, pp. 1989–2000, Jun. 2018.

[9] J. Cui, Y. Liu, Z. Ding, P. Fan, and A. Nallanathan, “Optimal user scheduling and power allocation for millimeter wave NOMA systems,” IEEE Trans. Wireless Commun., vol. 17, no. 3, pp. 1502–1517, Mar. 2018.

[10] Z. Xiao, L. Zhu, J. Choi, P. Xia, and X. G. Xia, “Joint power allocation and beamforming for non-orthogonal multiple access (NOMA) in 5G millimeter-wave communications,” IEEE Trans. Wireless Commun., vol. 17, no. 5, pp. 2961–2974, May 2019.

[11] L. Zhu, J. Zhang, Z. Xiao, X. Cao, D. O. Wu, and X. Xia, “Joint power control and beamforming for uplink non-orthogonal multiple access in 5G millimeter-wave communication,” IEEE Trans. Wireless Commun., vol. 17, no. 9, pp. 6177–6189, Sep. 2018.
[12] Z. Wei, L. Zhao, J. Guo, D. W. K. Ng, and J. Yuan, “A multi-beam NOMA framework for hybrid mmWave systems,” in Proc. IEEE Int. Conf. Commun. (ICC), Kansas City, MO, USA, May 2018, pp. 1–7.

[13] Z. Wei, L. Zhao, J. Guo, D. W. K. Ng, and J. Yuan, “Multi-beam NOMA for hybrid mmWave systems,” IEEE Trans. Commun., vol. 67, no. 2, pp. 1705–1719, Feb. 2019.

[14] Z. Ding, M. Peng, and H. V. Poor, “Cooperative non-orthogonal multiple access in 5G systems,” IEEE Commun. Lett., vol. 19, no. 8, pp. 1462–1465, Aug. 2015.

[15] Y. Zhou, W. S. Wong, and R. Schober, “Performance analysis of millimeter wave NOMA networks with beam misalignment,” in Proc. IEEE Int. Conf. Commun. (ICC), Kansas City, MO, USA, May 2018, pp. 1–7.

[16] W. Liang, Z. Ding, Y. Li, and L. Song, “User pairing for downlink non-orthogonal multiple access networks using matching algorithm,” IEEE Trans. Commun., vol. 65, no. 12, pp. 5319–5332, Dec. 2017.

[17] B. Di, L. Song, and Y. Li, “Sub-channel assignment, power allocation and user scheduling for non-orthogonal multiple access networks,” IEEE Trans. Wireless Commun., vol. 15, no. 11, pp. 7686–7698, Nov. 2016.

[18] Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, “Optimal joint power and subcarrier allocation for full-duplex multicarrier non-orthogonal multiple access systems,” IEEE Trans. Commun., vol. 65, no. 3, pp. 1077–1091, Mar. 2017.

[19] Z. Ding, P. Fan, and H. V. Poor, “Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions,” IEEE Trans. Veh. Technol., vol. 65, no. 8, pp. 6010–6023, Aug. 2016.

[20] Z. Ding, P. Fan, and H. V. Poor, “User pairing in non-orthogonal multiple access downlink transmissions,” in Proc. IEEE Global Commun. Conf. (GLOBECOM), San Diego, CA, USA, Dec. 2015, pp. 1–5.

[21] H. Zhang, D. Zhang, W. Meng, and C. Li, “User pairing algorithm with SIC in non-orthogonal multiple access system,” in Proc. IEEE Int. Conf. Commun. (ICC), Kuala Lumpur, Malaysia, May 2016, pp. 1–6.

[22] Y. Liu, Z. Ding, M. Elkashlan, and H. V. Poor, “Cooperative non-orthogonal multiple access with simultaneous wireless information and power transfer,” IEEE J. Sel. Areas Commun., vol. 34, no. 4, pp. 938–953, Apr. 2016.

[23] W. Yi, Y. Liu, and A. Nallanathan, “Exploiting multiple access in clustered millimeter wave networks: NOMA or OMA?” in Proc. IEEE Int. Conf. Commun. (ICC), Kansas City, MO, USA, May 2018, pp. 1–6.

[24] N. Rupasinghe, Y. Yapıcı, I. Güvenç, M. Ghosh, and Y. Kakishima, “Angle feedback for NOMA transmission in mmWave drone networks,” IEEE J. Sel. Topics Signal Process., vol. 13, no. 3, pp. 628–643, Jun. 2019.

[25] Y. Zhou, W. S. Wong, and R. Schober, “Coverage and rate analysis of millimeter wave NOMA networks with beam misalignment,” IEEE Trans. Wireless Commun., vol. 17, no. 12, pp. 8211–8227, Dec. 2018.

[26] L. Zhu, J. Zhang, Z. Xiao, X. Cao, D. O. Wu, and X. Xia, “Millimeter-wave NOMA with user grouping, power allocation and hybrid beamforming,” IEEE Trans. Wireless Commun., vol. 18, no. 11, pp. 5065–5079, Nov. 2019.

[27] J. Lee, G. Gil, and Y. H. Lee, “Exploiting spatial sparsity for estimating channels of hybrid MIMO systems in millimeter wave communications,” in Proc. IEEE Global Telecommun. Conf. (GLOBECOM), Austin, TX, USA, Dec. 2014, pp. 3326–3331.

[28] Z. Xiao, T. He, P. Xia, and X. Xia, “Hierarchical codebook design for beamforming training in millimeter wave communication,” IEEE Trans. Wireless Commun., vol. 15, no. 5, pp. 3380–3382, May 2016.

[29] X. Lin, S. Wu, C. Jiang, L. Kuang, J. Yan, and L. Hanzo, “Estimation of broadband multiuser millimeter wave massive MIMO-OFDM channels by exploiting their sparse structure,” IEEE Trans. Wireless Commun., vol. 17, no. 6, pp. 3959–3973, Jun. 2018.

[30] G. Lee, Y. Sung, and J. Seo, “Randomly-directional beamforming in millimeter-wave multiuser multiser downlink,” IEEE Trans. Wireless Commun., vol. 15, no. 2, pp. 1086–1100, Feb. 2016.

[31] H. L. Van Trees, Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory, vol. 1. Hoboken, NJ, USA: Wiley, 2002.