Coverage and Rate Analysis of Millimeter Wave NOMA Networks with Beam Misalignment

Yong Zhou, *Member, IEEE*, Vincent W.S. Wong, *Fellow, IEEE*, and Robert Schober, *Fellow, IEEE*

***Abstract*—Non-orthogonal multiple access (NOMA) and millimeter wave (mmWave) are two key enabling technologies for the fifth generation (5G) wireless networks. In this paper, we develop a general performance analysis framework for downlink NOMA transmission in mmWave networks with spatially random users taking into account link blockages and directional beamforming. To facilitate NOMA transmission in mmWave networks, we propose an angle-based user pairing strategy, where the base station first randomly selects one user and then pairs it with the line-of-sight user that has the minimum relative angle difference. The proposed strategy increases the probability that both NOMA users are covered by the main lobe created by directional beamforming. To account for the randomness of link blockages and user locations, we consider dynamic user ordering among the paired NOMA users. Tools from stochastic geometry are utilized to derive the coverage probability, outage sum rate, and ergodic sum rate of the proposed NOMA scheme, where beam misalignment at both the base station and users is taken into account. Simulations validate the performance analysis and show that the proposed NOMA scheme achieves a larger coverage probability and higher outage and ergodic sum rates than conventional NOMA with distance-based user pairing and orthogonal multiple access.**

***Index Terms*—Millimeter wave networks, non-orthogonal multiple access, beam misalignment, angle-based user pairing, coverage probability, ergodic sum rate.**

I. Introduction

Non-orthogonal multiple access (NOMA) [2]–[4] has the potential to significantly enhance the spectral efficiency and fairness of the fifth generation (5G) wireless networks, which is vital for meeting the rapidly increasing traffic demands. Different from orthogonal multiple access (OMA), with NOMA, a base station can simultaneously serve multiple users in the same frequency channel. For example, power-domain NOMA[1] employs superposition coding at the transmitter side and applies successive interference cancellation (SIC) at the receiver side to accomplish this.

NOMA has recently received considerable research interest [4]–[12]. System-level performance evaluations in [4] show that transmit power allocation and user pairing are two important design aspects for NOMA. The authors in [5] study the impact of user pairing on the outage probability and the sum rate of NOMA systems, where both fixed and cognitive radio based power allocation strategies are considered. The results in [5] show that the performance gain of NOMA over OMA is larger when users with diverse channel conditions are paired. To maximize the sum rate of the paired NOMA users, optimal power allocation schemes for multiple-input multiple-output (MIMO) and multicarrier NOMA systems are proposed in [6] and [7], respectively. The performance of downlink NOMA transmission under unsaturated traffic conditions is studied in [8], [9]. According to the NOMA decoding strategy, the user, which is allocated a lower transmit power, has to decode the signal intended for the other user before decoding its own signal. By exploiting this feature, a cooperative NOMA scheme is proposed in [10], where the lower power user helps forward the signal of the other user. To achieve both superimposed signal transmission and cooperative diversity in one time slot, the authors in [11], [12] propose a dynamic decode-and-forward based cooperative NOMA scheme. The aforementioned studies focus on resource allocation and performance analysis of NOMA for sub-6 GHz frequencies.

Millimeter wave (mmWave) is another key enabling technology for 5G networks [13]–[15]. Signals at mmWave frequencies (i.e., 30 – 300 GHz) are more sensitive to blockage effects compared to sub-6 GHz frequencies. The 3rd Generation Partnership Project (3GPP) reported stark differences in the propagation characteristics of line-of-sight (LOS) and non-line-of-sight (NLOS) links at mmWave frequencies [16]. Due to the short wavelength of mmWave signals, antenna arrays can be deployed at the base station and the devices to implement beamforming and to exploit array gains to ensure a sufficiently high received signal power. Based on real-world channel measurements, the authors in [17] evaluate the performance of mmWave networks. Their results show that the capacity achieved by mmWave cellular networks can be an order-of-magnitude higher than that of conventional

Manuscript received on Mar. 2, 2018; revised on Jul. 31, 2018 and Sep. 27, 2018; accepted on Oct. 1, 2018. This work was supported in part by the ShanghaiTech University start-up funds and the Natural Sciences and Engineering Research Council (NSERC) of Canada. This paper was presented in part at the *IEEE International Conference on Communications (ICC)*, Kansas, MO, May 2018 [1]. The editor coordinating the review of this paper and approving it for publication was Marco Di Renzo. *(Corresponding author: Vincent W.S. Wong.)*

Y. Zhou was with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada. He is now with the School of Information Science and Technology, ShanghaiTech University, Shanghai 201210, China. (e-mail: zhouyong@shanghaitech.edu.cn).

V. W.S. Wong is with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada (e-mail: vincentw@ece.ubc.ca).

R. Schober is with the Institute for Digital Communications, Friedrich-Alexander University of Erlangen-Nuremberg, Erlangen 91058, Germany (e-mail: robert.schober@fau.de).

[1]NOMA can be realized by exploiting different domains (e.g., power domain, code domain). In this paper, we focus on power-domain NOMA.

cellular networks. By using a sectored model to approximate the beamforming pattern, the authors in [18] analyze the coverage probability of mmWave networks for different link blockage models. Based on realistic channel and blockage models extracted from empirical data, the coverage probability and the data rate of mmWave cellular networks are derived in [19]. The authors in [20] propose sinc and cosine antenna pattern models to better approximate the actual beamforming pattern and derive coverage probabilities for both ad hoc and cellular networks. By considering the blockages caused by large vehicles, the authors in [21] derive the outage probability of highway vehicular communications, where mmWave base stations are deployed alongside the road. The authors in [22] provide a realistic-yet-tractable performance analysis by using a curve fitting tool to approximate the channel gains of the desired and interfering links. The authors in [23] study the impact of beam training and pilot reuse on the achievable rate of mmWave networks. When large antenna arrays are used, the authors in [24] compare the performance of different MIMO techniques (i.e., multi-user MIMO, single-user spatial multiplexing, and single-user analog beamforming) in mmWave networks.

The coexistence of NOMA and mmWave has recently been studied in [25]–[28]. The authors in [25] apply random beamforming in mmWave-NOMA networks to reduce the channel estimation overhead. Taking into account hardware constraints, the application of finite resolution analog beamforming in mmWave-NOMA networks is studied in [26]. The authors in [27] analyze the capacity of mmWave-NOMA networks in both noise-limited and interference-limited scenarios. To support more users than the number of available radio frequency (RF) chains, the authors in [28] propose a beamspace MIMO-NOMA scheme for mmWave networks. However, in practical implementations, beam misalignment at the base station and the users is inevitable [29]. This can reduce the probability that the paired NOMA users are covered by the main lobe created by directional beamforming and degrade the network performance. Hence, to overcome this problem, a new user pairing strategy and a corresponding performance analysis framework taking into account beam misalignment are needed to facilitate NOMA transmission in mmWave networks.

In this paper, we consider a downlink mmWave network with spatially random users. As the link blockage is distance-dependent, the spatial locations of the LOS and NLOS users are modeled as independent inhomogeneous Poisson point processes (PPPs). We take into account both directional beamforming and beam misalignment at both the base station and the paired NOMA users. To facilitate NOMA transmission in mmWave networks, we propose an angle-based user pairing strategy. Specifically, the base station first randomly selects one user and then pairs it with the LOS user that has the minimum relative angle difference. The proposed NOMA scheme increases the probability that both paired NOMA users are covered by the main lobe of the base station, even in the presence of beam misalignment. The main contributions of this paper are summarized as follows:

- We develop a general and tractable performance analysis framework for downlink NOMA transmission in mmWave networks with spatially random users taking into account link blockages, directional beamforming, beam misalignment, user pairing, and dynamic user ordering. We propose a novel angle-based user pairing strategy to facilitate NOMA transmission in mmWave networks by increasing the probability of exploiting antenna array gains.
- We derive the distributions of the random variables relevant for performance analysis, including the distances between the base station and the LOS (and NLOS) users, the angle difference between the paired NOMA users, and the total directivity gains between the base station and the paired NOMA users. Based on these results, we derive the coverage probability, the outage sum rate, and the ergodic sum rate of the proposed NOMA scheme, where Gauss-Chebyshev quadrature is utilized to reduce the computational complexity.
- Simulation results validate the performance analysis and demonstrate that the proposed NOMA scheme with angle-based user pairing achieves a better performance than conventional NOMA with distance-based user pairing and OMA in terms of the coverage probability, the outage sum rate, and the ergodic sum rate. Moreover, we illustrate the impact of various system parameters (e.g., beamsteering error variances, transmit power allocation coefficients, beamwidth of the main lobe, and user density) on the network performance.

This paper extends its conference version [1] in several aspects. Instead of fixed user ordering as in [1], we consider dynamic user ordering to account for the randomness of link blockages and user locations. Besides the coverage probability and the outage sum rate analyzed in [1], we also derive the ergodic sum rate of the proposed NOMA scheme, where adaptive modulation/coding is used to achieve the Shannon bound for a given instantaneous channel quality.

The remainder of this paper is organized as follows. In Section II, we describe the system model including the proposed angle-based user pairing and dynamic user ordering strategies. The coverage probability and the outage sum rate are analyzed in Section III, while the ergodic sum rate is derived in Section IV. Numerical and simulation results are presented in Section V. Finally, Section VI concludes this paper.

II. System Model

A. Network Topology

Consider downlink NOMA transmission in an mmWave cellular network, which consists of one base station and multiple spatially random users, as shown in Fig. 1. The base station is located at the center of a circular network coverage area with radius R. The spatial locations of the users are assumed to follow a homogeneous PPP, denoted as Φ, with density λ, which represents the average number of users per unit area.

As mmWave links are vulnerable to blockage effects, it is necessary to model both the LOS and NLOS path loss characteristics [30]. For outdoor transmission, a link can be either LOS or NLOS, depending on the presence of obstacles between the base station and the user. The blockages of different links are assumed to be independent from each other, as in [18]–[20], [24], [25]. Numerical results in [18], [31] show

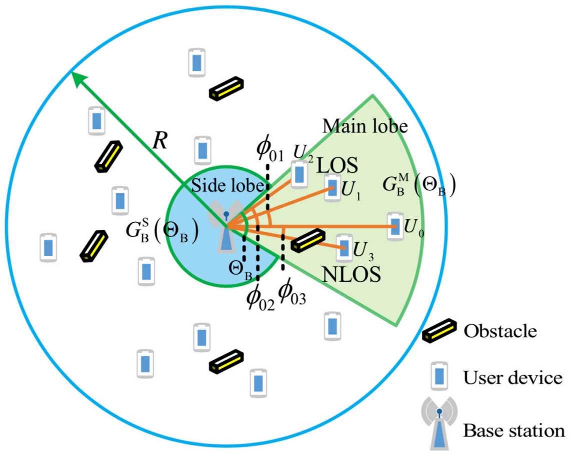


Fig. 1: Network topology for the considered downlink mmWave network. The base station observes both the LOS and NLOS users at mmWave frequencies. A sectored model is employed to approximate the beamforming pattern. The angle difference between users U_0 and U_i is denoted as ϕ_{0i}.

that ignoring the correlation of blockage effects usually causes only a minor loss in accuracy of the signal-to-interference-plus-noise ratio (SINR). The probability that a link of length r is LOS is $p(r) = \exp(-r/\eta)$, as in [18], [31], [32], where η is the average LOS range. The value of η depends on the average size and density of obstacles. Based on the channel measurements in [17], the path losses of the LOS and NLOS links can be expressed as $\ell_{\mathrm{L}}(r) = C_{\mathrm{L}} r^{-\beta_{\mathrm{L}}}$ and $\ell_{\mathrm{N}}(r) = C_{\mathrm{N}} r^{-\beta_{\mathrm{N}}}$, respectively, where C_{L} and C_{N} are constant parameters, and β_{L} and β_{N} are the path loss exponents of the LOS and NLOS links, respectively. Thus, the path loss of the link between the base station and user U_i is

$$\ell(r_i) = \mathbb{B}(p(r_i))\ell_{\mathrm{L}}(r_i) + (1 - \mathbb{B}(p(r_i)))\ell_{\mathrm{N}}(r_i), \tag{1}$$

where $\mathbb{B}(x)$ is a Bernoulli random variable with mean x, and r_i denotes the Euclidean distance between the base station and user U_i. The blockage of a link is distance-dependent, and thus, both the LOS and NLOS users are not homogeneously distributed. In particular, the homogeneous PPP Φ can be divided into two independent inhomogeneous PPPs [33], [34], denoted as Φ_{L} and $\Phi_{\mathrm{N}} = \Phi \backslash \Phi_{\mathrm{L}}$, which comprise the locations of the LOS and NLOS users, respectively. The densities of inhomogeneous PPPs Φ_{L} and Φ_{N} are given by $\lambda p(r)$ and $\lambda(1 - p(r))$, respectively.

Due to the limited scattering at mmWave frequencies and the adoption of directional beamforming, each link is assumed to suffer from independent quasi-static Nakagami-m fading, as in [18], [24], [25], [35]. The parameters of the Nakagami-m fading for the LOS and NLOS links are denoted as N_{L} and N_{N}, respectively, and are assumed to be positive integers for simplicity. As a result, the channel fading coefficient between the base station and user U_i, denoted as $|h_i|^2$, is a normalized Gamma random variable, i.e., $\mathrm{Gamma}(N_\nu, 1/N_\nu), \nu \in \{\mathrm{L}, \mathrm{N}\}$. As the delay spread is usually small at mmWave frequencies, frequency selective fading is ignored. Furthermore, since the channel coherence time is inversely proportional to the carrier frequency, for a given mobile velocity, mmWave channels change faster than sub-6 GHz channels [14]. However, fast-changing channels make successful SIC decoding challenging and complicate the power allocation for the paired NOMA users. Hence, in this paper, we assume a low-mobility scenario to facilitate SIC decoding.

B. Directional Beamforming

Due to the short wavelength of mmWave signals, antenna arrays are deployed at both the base station and the users to perform directional beamforming, which can help overcome the high noise power introduced by the large spectrum bandwidth. To make the performance analysis of the proposed NOMA scheme tractable, we adopt the sectored model[2] to approximate the actual beamforming pattern, as shown in Fig. 1. This model has also been adopted in other recent studies, see [18], [23], [32]. The directional antenna gain is given by

$$G_{\mathrm{a}}(\Theta_{\mathrm{a}}, \phi) = \begin{cases} G_{\mathrm{a}}^{\mathrm{M}}(\Theta_{\mathrm{a}}) = \frac{2\pi}{\Theta_{\mathrm{a}}} \frac{\gamma_{\mathrm{a}}}{\gamma_{\mathrm{a}}+1}, & \text{if } |\phi| \le \frac{1}{2}\Theta_{\mathrm{a}}, \\ G_{\mathrm{a}}^{\mathrm{S}}(\Theta_{\mathrm{a}}) = \frac{2\pi}{2\pi - \Theta_{\mathrm{a}}} \frac{1}{\gamma_{\mathrm{a}}+1}, & \text{if } |\phi| > \frac{1}{2}\Theta_{\mathrm{a}}, \end{cases} \tag{2}$$

where $G_{\mathrm{a}}^{\mathrm{M}}(\Theta_{\mathrm{a}})$ and $G_{\mathrm{a}}^{\mathrm{S}}(\Theta_{\mathrm{a}})$ denote the antenna gains of the main lobe and the side lobe, respectively, Θ_{a} denotes the beamwidth of the main lobe, $\mathrm{a} \in \{\mathrm{B}, \mathrm{U}\}$ represents either the base station (B) or a user (U), ϕ denotes the angle off the boresight direction, $\gamma_{\mathrm{a}} = \frac{2\pi}{\varsigma(2\pi - \Theta_{\mathrm{a}})}$ is the forward-to-backward power ratio, and ς is a constant. The sectored-pattern antenna model in (2) ensures that the total transmit power is constant, i.e., $\int_0^{2\pi} G_{\mathrm{a}}(\Theta_{\mathrm{a}}, \phi) \mathrm{d}\phi = G_{\mathrm{a}}^{\mathrm{M}}(\Theta_{\mathrm{a}}) \frac{\Theta_{\mathrm{a}}}{2\pi} + G_{\mathrm{a}}^{\mathrm{S}}(\Theta_{\mathrm{a}}) \frac{2\pi - \Theta_{\mathrm{a}}}{2\pi} = 1$. There exists a tradeoff between the beamwidth and the antenna gain of the main lobe. A wider beamwidth of the main lobe leads to a smaller antenna gain of the main lobe, and vice versa.

The base station and the users can adjust their beam orientations to cover their intended receiver/transmitter in the main lobe. However, beam misalignment is inevitable in practical implementations [29]. We denote the additive beamsteering errors of the base station and user U_i as Δ_{B} and Δ_{U_i}, respectively. We assume that Δ_{B} and $\{\Delta_{\mathrm{U}_i}, z_i \in \Phi\}$ are independent and identically distributed, where z_i denotes the spatial location of user U_i. The beamsteering errors are assumed to follow a Gaussian distribution with zero mean and variance $\sigma_{\mathrm{a}}^2, \mathrm{a} \in \{\mathrm{B}, \mathrm{U}\}$. Hence, the cumulative distribution functions (CDFs) of the beamsteering errors at the base station and the users can be expressed as $F_{\Delta_{\mathrm{B}}}(x) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{x}{\sigma_{\mathrm{B}}\sqrt{2}}\right)\right)$ and $F_{\Delta_{\mathrm{U}}}(x) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{x}{\sigma_{\mathrm{U}}\sqrt{2}}\right)\right)$, respectively, where $\mathrm{erf}(\cdot)$ denotes the error function.

[2]Although simplified, the sectored model captures the key features of the actual beamforming pattern, including the directivity gain, the front-to-back ratio, and the half-power beamwidth [18]. An even more accurate model may better characterize the network performance [20]. However, it would also lead to a more complicated analysis and numerical evaluation. The performance analysis framework developed in this paper can be extended to incorporate more accurate beamforming pattern models by taking into account the corresponding array gain. However, the analysis would become very complicated if the user pairing and beam misalignment were also taken into account, which is the main focus of this paper. Therefore and due to space limitation, the analysis for more accurate beamforming pattern models will be considered in our future work.

C. NOMA Scheme for mmWave Networks

We consider the case where two users are paired for NOMA transmission. A two-user NOMA scheme has recently been included in the 3GPP Long Term Evolution Advanced (LTE-A) standard [36], where it is referred to as multiuser superposition transmission (MUST). To ensure fairness, the base station first selects one user at random. This user is referred to as the *typical user* and denoted as U_0. To facilitate NOMA transmission in mmWave networks with directional beamforming, we propose an angle-based user pairing strategy. Specifically, the user paired with typical user U_0 is denoted as U_{p} and is selected based on the following criterion

$$z_{\mathrm{p}} = \begin{cases} \arg \min\limits_{z_i \in \Phi_{\mathrm{L}} \setminus \{z_0\}} \{\phi_{0i}\}, & \text{if user } U_0 \text{ is LOS}, \\ \arg \min\limits_{z_i \in \Phi_{\mathrm{L}}} \{\phi_{0i}\}, & \text{if user } U_0 \text{ is NLOS}, \end{cases} \tag{3}$$

where z_{p} denotes the spatial location of the paired NOMA user U_{p}, $\phi_{0i} = |\phi_0 - \phi_i|$ denotes the absolute value of the angle difference between angles ϕ_0 and ϕ_i, and ϕ_i denotes the angle of user U_i with respect to the base station. The proposed user pairing strategy selects the LOS user that has the minimum angle difference (i.e., $\phi_{0\mathrm{p}}$) to the typical user. Hence, the angle information of each LOS user has to be known at the base station. For the example shown in Fig. 1, user U_1 is selected as the paired NOMA user for the typical user, because user U_3 is NLOS and $\phi_{01} < \phi_{02}$. We note that for the proposed angle-based user pairing strategy, the probability that both NOMA users are covered by the main lobe of the base station is high, even for narrow beamwidths. Hence, the probability that the paired NOMA users can exploit the antenna array gain to enhance the signal quality is also high. On the other hand, the conventional distance-based user pairing strategy was developed for omni-directional transmission [37]. Since the angle difference is not taken into account, it is not likely that the paired NOMA users are both covered by the main lobe created by directional beamforming. Hence, they cannot both exploit the antenna array gain, which may lead to a large performance degradation in mmWave networks.

Whether NOMA or OMA is enabled depends on the value of angle difference $\phi_{0\mathrm{p}}$. In particular, the base station performs NOMA and simultaneously serves both NOMA users (i.e., U_0 and U_{p}) if $\phi_{0\mathrm{p}} \leq \Theta_{\mathrm{B}}^{\mathrm{NOMA}}$, where $\Theta_{\mathrm{B}}^{\mathrm{NOMA}}$ denotes the beamwidth of the main lobe of the base station for NOMA transmission. On the other hand, if $\phi_{0\mathrm{p}} > \Theta_{\mathrm{B}}^{\mathrm{NOMA}}$, i.e., the base station cannot cover both NOMA users in the main lobe even without beam misalignment, then the base station serves users U_0 and U_{p} using OMA with beamwidth $\Theta_{\mathrm{B}}^{\mathrm{OMA}}$. The beamwidth of the main lobe at the base station for NOMA transmission is generally not smaller than that for OMA transmission, i.e., $\Theta_{\mathrm{B}}^{\mathrm{NOMA}} \geq \Theta_{\mathrm{B}}^{\mathrm{OMA}}$. This is because although a narrow beam may be easily steered to cover a single user, it may not be able to simultaneously cover two paired NOMA users.

For NOMA transmission (i.e., $\phi_{0\mathrm{p}} \leq \Theta_{\mathrm{B}}^{\mathrm{NOMA}}$), the signal intended for the typical user can either be decoded first or second, depending on the link blockages and the spatial locations of users U_0 and U_{p}. In particular, the signal intended for paired NOMA user U_{p} is decoded first when both NOMA users are LOS and $r_0 \leq r_{\mathrm{p}}$. The corresponding NOMA scheme is referred to as NOMA-I. On the other hand, the signal intended for typical user U_0 is decoded first when one of the following cases occurs: (1) Both NOMA users are LOS and $r_{\mathrm{p}} < r_0$; (2) The typical user is NLOS and the paired NOMA user is LOS. The NOMA schemes for cases (1) and (2) are referred to as NOMA-II and NOMA-III, respectively. We denote the transmit power allocation coefficients for the signals being decoded first and second as α_{F} and α_{S}, respectively, with $\alpha_{\mathrm{F}} > \alpha_{\mathrm{S}}$ and $\alpha_{\mathrm{F}}^2 + \alpha_{\mathrm{S}}^2 = 1$. Note that we consider a fixed transmit power allocation strategy, where the values of α_{F} and α_{S} are pre-allocated and do not depend on the channel gains. The signal reception for each case is described as follows.

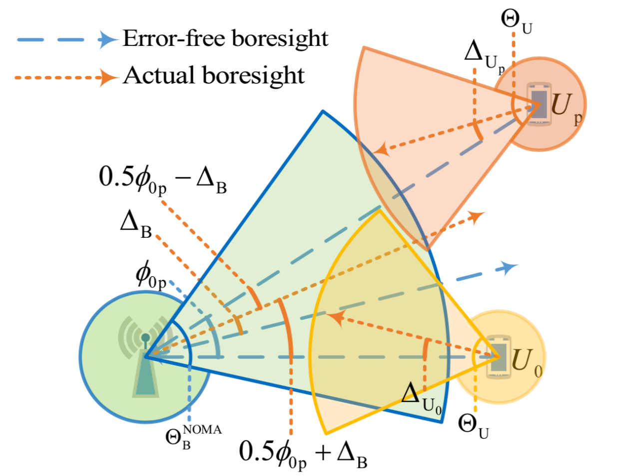


Fig. 2: Directional beamforming at the base station and at NOMA users U_0 and U_{p} with beamsteering errors Δ_{B}, $\Delta_{\mathrm{U_0}}$, and $\Delta_{\mathrm{U_p}}$. The error-free boresight direction of the base station has the same angle difference (i.e., $0.5\phi_{0\mathrm{p}}$) to typical user U_0 and paired NOMA user U_{p}. However, due to beam misalignment, there is in general a beamsteering error (i.e., Δ_{B}) between the actual and the error-free boresight directions.

1) NOMA-I: In this case, the signal intended for user U_{p} is decoded first. The SINR of the signal intended for user U_{p} observed at user U_0 is

$$\gamma_{\mathrm{p}\to 0} = \frac{\alpha_{\mathrm{F}}^2 P_{\mathrm{B}} |h_0|^2 \mathcal{D}_0^{\mathrm{NOMA}} \ell_{\mathrm{L}}(r_0)}{\alpha_{\mathrm{S}}^2 P_{\mathrm{B}} |h_0|^2 \mathcal{D}_0^{\mathrm{NOMA}} \ell_{\mathrm{L}}(r_0) + \sigma^2}, \tag{4}$$

where P_{B} denotes the total transmit power of the base station, $\mathcal{D}_0^{\mathrm{NOMA}}$ denotes the total directivity gain between the base station and typical user U_0 when NOMA is enabled, and σ^2 denotes the power of the additive white Gaussian noise (AWGN).

To reduce the adverse effects of beam misalignment on the coverage of the paired NOMA users in the main lobe, the base station adjusts its beam orientation so that its error-free boresight direction has the same angle difference (i.e., $0.5\phi_{0\mathrm{p}}$) with respect to typical user U_0 and paired NOMA user U_{p}, as shown in Fig. 2. However, due to beam misalignment, there is a beamsteering error (i.e., Δ_{B}) between the actual and the error-free boresight directions. Hence, for NOMA, the total directivity gain between the base station and typical user U_0 is given by

$$\mathcal{D}_0^{\mathrm{NOMA}} = G_{\mathrm{B}}\left(\Theta_{\mathrm{B}}^{\mathrm{NOMA}}, 0.5\phi_{0\mathrm{p}} + \Delta_{\mathrm{B}}\right) G_{\mathrm{U}}\left(\Theta_{\mathrm{U}}, \Delta_{\mathrm{U_0}}\right), \tag{5}$$

where Θ_{U} denotes the beamwidth of the main lobe of the user.

If user U_0 can successfully decode the signal intended for user U_{p}, then user U_0 performs SIC to remove intra-cell interference and decodes its own signal with signal-to-noise ratio (SNR)

$$\gamma_0 = \frac{\alpha_{\mathrm{S}}^2 P_{\mathrm{B}} \left|h_0\right|^2 \mathcal{D}_0^{\mathrm{NOMA}} \ell_{\mathrm{L}}(r_0)}{\sigma^2}. \tag{6}$$

Otherwise, user U_0 cannot perform SIC and cannot decode its own signal, i.e., an outage occurs.

On the other hand, paired NOMA user U_{p} treats the signal intended for user U_0 as intra-cell interference. Hence, the SINR observed at the paired NOMA user U_{p} is given by

$$\gamma_{\mathrm{p}|0} = \frac{\alpha_{\mathrm{F}}^2 P_{\mathrm{B}} \left|h_{\mathrm{p}}\right|^2 \mathcal{D}_{\mathrm{p}}^{\mathrm{NOMA}} \ell_{\mathrm{L}}(r_{\mathrm{p}})}{\alpha_{\mathrm{S}}^2 P_{\mathrm{B}} \left|h_{\mathrm{p}}\right|^2 \mathcal{D}_{\mathrm{p}}^{\mathrm{NOMA}} \ell_{\mathrm{L}}(r_{\mathrm{p}}) + \sigma^2}, \tag{7}$$

where $\mathcal{D}_{\mathrm{p}}^{\mathrm{NOMA}}$ denotes the total directivity gain between the base station and paired NOMA user U_{p} when NOMA is enabled. Similar to (5), we have

$$\mathcal{D}_{\mathrm{p}}^{\mathrm{NOMA}} = G_{\mathrm{B}}\left(\Theta_{\mathrm{B}}^{\mathrm{NOMA}}, 0.5\phi_{0\mathrm{p}} - \Delta_{\mathrm{B}}\right) G_{\mathrm{U}}\left(\Theta_{\mathrm{U}}, \Delta_{\mathrm{U}_{\mathrm{p}}}\right), \tag{8}$$

where beamsteering error $\Delta_{U_{\mathrm{p}}}$ is independent of beamsteering errors Δ_{B} and Δ_{U_0}.

2) NOMA-II and NOMA-III: In these two cases, the signal intended for user U_0 is decoded first. Paired NOMA user U_{p} first decodes the signal intended for typical user U_0 with SINR

$$\gamma_{0\to\mathrm{p}} = \frac{\alpha_{\mathrm{F}}^2 P_{\mathrm{B}} \left|h_{\mathrm{p}}\right|^2 \mathcal{D}_{\mathrm{p}}^{\mathrm{NOMA}} \ell_{\mathrm{L}}(r_{\mathrm{p}})}{\alpha_{\mathrm{S}}^2 P_{\mathrm{B}} \left|h_{\mathrm{p}}\right|^2 \mathcal{D}_{\mathrm{p}}^{\mathrm{NOMA}} \ell_{\mathrm{L}}(r_{\mathrm{p}}) + \sigma^2}. \tag{9}$$

If paired NOMA user U_{p} can successfully decode the signal intended for typical user U_0, then user U_{p} performs SIC and decodes its own signal with the following SNR

$$\gamma_{\mathrm{p}} = \frac{\alpha_{\mathrm{S}}^2 P_{\mathrm{B}} \left|h_{\mathrm{p}}\right|^2 \mathcal{D}_{\mathrm{p}}^{\mathrm{NOMA}} \ell_{\mathrm{L}}(r_{\mathrm{p}})}{\sigma^2}. \tag{10}$$

On the other hand, typical user U_0 decodes its own signal based on the following SINR

$$\gamma_{0|\mathrm{p}} = \frac{\alpha_{\mathrm{F}}^2 P_{\mathrm{B}} \left|h_0\right|^2 \mathcal{D}_0^{\mathrm{NOMA}} \ell_{\nu}(r_0)}{\alpha_{\mathrm{S}}^2 P_{\mathrm{B}} \left|h_0\right|^2 \mathcal{D}_0^{\mathrm{NOMA}} \ell_{\nu}(r_0) + \sigma^2}, \tag{11}$$

where $\nu \in \{\mathrm{L}, \mathrm{N}\}$ represents either the LOS (L) or the NLOS (N) link.

3) OMA: When a second NOMA user for pairing does not exist or the angle difference $\phi_{0\mathrm{p}}$ is greater than beamwidth $\Theta_{\mathrm{B}}^{\mathrm{NOMA}}$, the base station cannot cover two NOMA users in the main lobe, even without beam misalignment. Thereby, the base station serves users U_0 and U_{p} using OMA with main lobe beamwidth $\Theta_{\mathrm{B}}^{\mathrm{OMA}}$. The received SNR observed at user $U_j, j \in \{0, \mathrm{p}\}$, is

$$\gamma_{j,\nu}^{\mathrm{OMA}} = \frac{P_{\mathrm{B}} \left|h_j\right|^2 \mathcal{D}_j^{\mathrm{OMA}} \ell_{\nu}(r_j)}{\sigma^2}, \quad \nu \in \{\mathrm{L}, \mathrm{N}\}, \tag{12}$$

where $\mathcal{D}_j^{\mathrm{OMA}}$ denotes the total directivity gain between the base station and user U_j when OMA is enabled. For OMA transmission, the base station adjusts its beam orientation so that its error-free boresight direction is aligned with the vector from the base station to user U_j. Considering the beamsteering errors at both the base station and user U_j, the total directivity gain between the base station and user $U_j, j \in \{0, \mathrm{p}\}$, for OMA is given by

$$\mathcal{D}_j^{\mathrm{OMA}} = G_{\mathrm{B}}\left(\Theta_{\mathrm{B}}^{\mathrm{OMA}}, \Delta_{\mathrm{B}}\right) G_{\mathrm{U}}\left(\Theta_{\mathrm{U}}, \Delta_{\mathrm{U}_j}\right). \tag{13}$$

The SINRs and the total directivity gains given in (4)-(13) are all random variables, due to the randomness of the small-scale fading (h_0 and h_{p}), the angle difference ($\phi_{0\mathrm{p}}$), the beamsteering errors (Δ_{B}, Δ_{U_0}, and $\Delta_{\mathrm{U}_{\mathrm{p}}}$), the distance between the base station and the NOMA users (r_0 and r_{p}), and the link blockages. In the following two sections, we analyze the coverage probability, the outage sum rate, and the ergodic sum rate of the proposed NOMA scheme.

III. Coverage Probability and Outage Sum Rate Analysis

In this section, we first present the probability density functions (PDFs) and probability mass functions (PMFs) of the relevant random variables, and then we analyze the coverage probabilities and outage sum rate of the paired users for the proposed NOMA scheme.

A. Some Useful PDFs and PMFs

To derive the coverage probability, we first need to derive the PDFs and PMFs of the relevant random variables, including the distances between the base station and the users, the number of LOS and NLOS users in the network, the angle difference between the paired NOMA users, and the total directivity gains between the base station and the users. The following lemma presents the PDFs of the distances of the randomly selected LOS and NLOS users to the base station.

Lemma 1. *Given that the base station observes at least one LOS user, the PDF of the distance between the base station and a randomly selected LOS user is given by*

$$f_{\mathrm{L}}(r) = \frac{r}{\rho} \exp\left(-\frac{r}{\eta}\right), \quad 0 \le r \le R, \tag{14}$$

where $\rho = \eta^2 \left(1 - \exp\left(-\frac{R}{\eta}\right)\left(1 + \frac{R}{\eta}\right)\right)$.

Similarly, given that the base station observes at least one NLOS user, the PDF of the distance between the base station and a randomly selected NLOS user can be expressed as

$$f_{\mathrm{N}}(r) = \frac{2r}{R^2 - 2\rho}\left(1 - \exp\left(-\frac{r}{\eta}\right)\right), \quad 0 \le r \le R. \tag{15}$$

Proof. Please refer to Appendix A. □

For two LOS users, we refer to the LOS user closer to the base station as the *near LOS user* and to the other LOS user as the *far LOS user*. The following lemma presents the PDFs of the distances of the near and far LOS users from the base station.

Lemma 2. *For two LOS users, the PDF of the distance between the base station and the near LOS user is given by*

$$f_{r_{\mathrm{n}}}(r) = \frac{2r}{\rho}\left(1 - \frac{\eta^2}{\rho}\right) \exp\left(-\frac{r}{\eta}\right)$$

$$+\frac{2\eta^2 r}{\rho^2}\left(1+\frac{r}{\eta}\right)\exp\left(-\frac{2r}{\eta}\right), 0\le r\le R. \quad (16)$$

The PDF of the distance between the base station and the far LOS user can be expressed as

$$f_{r_{\rm f}}(r)=\frac{2\eta^2 r}{\rho^2}\exp\left(-\frac{r}{\eta}\right) - \frac{2\eta^2 r}{\rho^2}\left(1+\frac{r}{\eta}\right)\exp\left(-\frac{2r}{\eta}\right), 0\le r\le R. \quad (17)$$

Proof. Please refer to Appendix B. □

According to Lemmas 1 and 2, the PDFs of the user distances only depend on the average LOS range (η) and the radius of the network coverage area (R). As the spatial locations of the LOS and NLOS users form inhomogeneous PPPs $\Phi_{\rm L}$ and $\Phi_{\rm N}$ with densities $\lambda p(r)$ and $\lambda(1-p(r))$, respectively, the PMFs of the number of the LOS and NLOS users in the network coverage area can be expressed as $\Psi_{\rm L}(k_{\rm L})=(\lambda 2\pi\rho)^{k_{\rm L}}\exp(-\lambda 2\pi\rho)/k_{\rm L}!$ and $\Psi_{\rm N}(k_{\rm N})=\left(\lambda\pi\left(R^2-2\rho\right)\right)^{k_{\rm N}}\exp\left(-\lambda\pi\left(R^2-2\rho\right)\right)/k_{\rm N}!$, respectively. When the number of LOS users in the network coverage area is given, the following lemma provides the PDF of angle difference $\phi_{0\rm p}$.

Lemma 3. *When there are $k_{\rm L}$ LOS users in the network coverage area, based on the user pairing strategy given in (3), the PDF of the angle difference between the typical user and the paired NOMA user, i.e., $\phi_{0\rm p}$, can be expressed as*

$$f_{\phi_{0\rm p}}(\phi)=\begin{cases}\frac{k_{\rm L}-1}{\pi}\left(\frac{2\pi-\phi}{2\pi}\right)^{2k_{\rm L}-3}, & \text{if } U_0 \text{ is LOS},\\ \frac{k_{\rm L}}{\pi}\left(\frac{2\pi-\phi}{2\pi}\right)^{2k_{\rm L}-1}, & \text{if } U_0 \text{ is NLOS}.\end{cases} \quad (18)$$

Proof. Please refer to Appendix C. □

According to Lemma 3, the PDF of angle difference $\phi_{0\rm p}$ is independent of the user locations, but depends on whether typical user U_0 is LOS or NLOS. From (18), we observe that the probability of having a small angle difference ($\phi_{0\rm p}$) increases with $k_{\rm L}$, which directly depends on the user density (λ). A smaller angle difference increases the probability that NOMA transmission is enabled as well as the probability that both NOMA users are covered by the main lobe in the presence of beam misalignment. Given the angle difference $\phi_{0\rm p}$, the following lemma presents the PMFs of the total directivity gains $\mathcal{D}_0^{\rm NOMA}$ and $\mathcal{D}_{\rm p}^{\rm NOMA}$ for NOMA transmission.

Lemma 4. *Given the angle difference between typical user U_0 and paired NOMA user $U_{\rm p}$, i.e., $\phi_{0\rm p}$, the PMFs of the total directivity gains $\mathcal{D}_0^{\rm NOMA}$ and $\mathcal{D}_{\rm p}^{\rm NOMA}$ for NOMA transmission can, respectively, be expressed as*

$$\mathbb{P}\left(\mathcal{D}_0^{\rm NOMA}=c_i\right)=d_i\left(\phi_{0\rm p}\right), \quad \text{for } i=\{1,2,3,4\}, \quad (19)$$
$$\mathbb{P}\left(\mathcal{D}_{\rm p}^{\rm NOMA}=c_i\right)=v_i\left(\phi_{0\rm p}\right), \quad \text{for } i=\{1,2,3,4\}, \quad (20)$$

where

$$\begin{aligned}&c_1=G_{\rm B}^{\rm M}\left(\Theta_{\rm B}^{\rm NOMA}\right)G_{\rm U}^{\rm M}\left(\Theta_{\rm U}\right), c_2=G_{\rm B}^{\rm M}\left(\Theta_{\rm B}^{\rm NOMA}\right)G_{\rm U}^{\rm S}\left(\Theta_{\rm U}\right),\\&c_3=G_{\rm B}^{\rm S}\left(\Theta_{\rm B}^{\rm NOMA}\right)G_{\rm U}^{\rm M}\left(\Theta_{\rm U}\right),\\&c_4=G_{\rm B}^{\rm S}\left(\Theta_{\rm B}^{\rm NOMA}\right)G_{\rm U}^{\rm S}\left(\Theta_{\rm U}\right),\end{aligned} \quad (21)$$

$$\begin{aligned}&d_1(\phi_{0\rm p})=g\left(\phi_{0\rm p}\right)y, \quad d_2\left(\phi_{0\rm p}\right)=g\left(\phi_{0\rm p}\right)(1-y),\\&d_3(\phi_{0\rm p})=\left(1-g\left(\phi_{0\rm p}\right)\right)y,\\&d_4\left(\phi_{0\rm p}\right)=\left(1-g\left(\phi_{0\rm p}\right)\right)(1-y),\end{aligned} \quad (22)$$

$$\begin{aligned}&v_1(\phi_{0\rm p})=s\left(\phi_{0\rm p}\right)y, \quad v_2(\phi_{0\rm p})=s\left(\phi_{0\rm p}\right)(1-y),\\&v_3(\phi_{0\rm p})=\left(1-s\left(\phi_{0\rm p}\right)\right)y,\\&v_4(\phi_{0\rm p})=\left(1-s\left(\phi_{0\rm p}\right)\right)(1-y),\end{aligned} \quad (23)$$

$g\left(\phi_{0\rm p}\right)=\frac{1}{2}\left(\operatorname{erf}\left(\frac{\Theta_{\rm B}^{\rm NOMA}-\phi_{0\rm p}}{2\sqrt{2}\sigma_{\rm B}}\right)-\operatorname{erf}\left(-\frac{\Theta_{\rm B}^{\rm NOMA}+\phi_{0\rm p}}{2\sqrt{2}\sigma_{\rm B}}\right)\right)$, $y=\frac{1}{2}\left(\operatorname{erf}\left(\frac{\Theta_{\rm U}}{2\sqrt{2}\sigma_{\rm U}}\right)-\operatorname{erf}\left(-\frac{\Theta_{\rm U}}{2\sqrt{2}\sigma_{\rm U}}\right)\right)$, *and* $s\left(\phi_{0\rm p}\right)=\frac{1}{2}\left(\operatorname{erf}\left(\frac{\Theta_{\rm B}^{\rm NOMA}+\phi_{0\rm p}}{2\sqrt{2}\sigma_{\rm B}}\right)-\operatorname{erf}\left(-\frac{\Theta_{\rm B}^{\rm NOMA}-\phi_{0\rm p}}{2\sqrt{2}\sigma_{\rm B}}\right)\right)$.

Proof. Please refer to Appendix D. □

According to Lemma 4, the total directivity gain depends on the beamwidth of the main lobes of the base station and the user, the beamsteering error variances, and the angle difference $\phi_{0\rm p}$. A larger beamwidth of the main lobe and a smaller angle difference reduce the negative impact of beam misalignment on the total directivity gain. Similarly, the following lemma presents the PMF of the total directivity gain $\mathcal{D}_0^{\rm OMA}$ for OMA transmission.

Lemma 5. *The PMF of the total directivity gain $\mathcal{D}_j^{\rm OMA}, j\in\{0,{\rm p}\}$, for OMA transmission is*

$$\mathbb{P}\left(\mathcal{D}_j^{\rm OMA}=l_i\right)=\tau_i, \quad \text{for } i=\{1,2,3,4\}, \quad (24)$$

where

$$\begin{aligned}&l_1=G_{\rm B}^{\rm M}\left(\Theta_{\rm B}^{\rm OMA}\right)G_{\rm U}^{\rm M}\left(\Theta_{\rm U}\right), && l_2=G_{\rm B}^{\rm M}\left(\Theta_{\rm B}^{\rm OMA}\right)G_{\rm U}^{\rm S}\left(\Theta_{\rm U}\right),\\&l_3=G_{\rm B}^{\rm S}\left(\Theta_{\rm B}^{\rm OMA}\right)G_{\rm U}^{\rm M}\left(\Theta_{\rm U}\right), && l_4=G_{\rm B}^{\rm S}\left(\Theta_{\rm B}^{\rm OMA}\right)G_{\rm U}^{\rm S}\left(\Theta_{\rm U}\right),\\&\tau_1=wy, && \tau_2=w(1-y),\\&\tau_3=(1-w)y, && \tau_4=(1-w)(1-y),\end{aligned} \quad (25)$$

$w=\frac{1}{2}\left(\operatorname{erf}\left(\frac{\Theta_{\rm B}^{\rm OMA}}{2\sqrt{2}\sigma_{\rm B}}\right)-\operatorname{erf}\left(-\frac{\Theta_{\rm B}^{\rm OMA}}{2\sqrt{2}\sigma_{\rm B}}\right)\right)$, *and y is given in Lemma 4.*

Proof. Please refer to Appendix E. □

Based on the results provided in Lemmas 1-5, we derive the coverage probabilities of typical user U_0 and paired NOMA user $U_{\rm p}$ in the following two subsections.

B. Coverage Probability of NOMA Transmission

Because of the dynamic user ordering, we denote the target data rate of the user whose signal is decoded first as $R_{\rm F}$ and the target data rate of the other user as $R_{\rm S}$. The coverage probability is defined as the probability that a user can successfully decode the signal transmitted by the base station with a certain target data rate. The signal intended for the typical user can either be decoded first or second depending on the link blockage status of the typical user and the distances of users U_0 and $U_{\rm p}$. Hence, we have the following three cases.

$$Q_{\mathrm{I},0}(\xi_1)=\sum_{n=0}^{N_{\mathrm{L}}-1}\sum_{i=1}^{4}\frac{(N_{\mathrm{L}}\xi_1)^n}{n!\,(c_iC_{\mathrm{L}})^n}\int_0^{\Theta_{\mathrm{B}}^{\mathrm{NOMA}}}d_i(\phi)f_{\phi_{0\mathrm{p}}}(\phi)\mathrm{d}\phi\int_0^{R}r^{n\beta_{\mathrm{L}}}\exp\left(-\frac{N_{\mathrm{L}}\xi_1r^{\beta_{\mathrm{L}}}}{c_iC_{\mathrm{L}}}\right)f_{r_{\mathrm{n}}}(r)\mathrm{d}r \tag{27}$$

$$\approx\sum_{n=0}^{N_{\mathrm{L}}-1}\sum_{i=1}^{4}\frac{\Theta_{\mathrm{B}}^{\mathrm{NOMA}}R\pi^2\,(N_{\mathrm{L}}\xi_1)^n}{4M_1M_2n!\,(c_iC_{\mathrm{L}})^n}\left(\sum_{m_1=1}^{M_1}\sqrt{1-\zeta_{m_1}^2}d_i\left(\phi_{0\mathrm{p},m_1}\right)f_{\phi_{0\mathrm{p}}}\left(\phi_{0\mathrm{p},m_1}\right)\right)$$

$$\times\left(\sum_{m_2=1}^{M_2}\sqrt{1-\varpi_{m_2}^2}r_{\mathrm{n},m_2}^{n\beta_{\mathrm{L}}}\exp\left(-\frac{N_{\mathrm{L}}\xi_1r_{\mathrm{n},m_2}^{\beta_{\mathrm{L}}}}{c_iC_{\mathrm{L}}}\right)f_{r_{\mathrm{n}}}\left(r_{\mathrm{n},m_2}\right)\right), \tag{28}$$

1) NOMA-I: In this case, typical user U_0 is the near LOS user. Hence, user U_0 can successfully decode its own signal if it can successfully perform SIC (i.e., $\gamma_{\mathrm{p}\rightarrow 0}\geq T_{\mathrm{F}}$) and the SNR of its own signal is not smaller than T_{S} (i.e., $\gamma_0\geq T_{\mathrm{S}}$), where $T_{\mathrm{F}}=2^{R_{\mathrm{F}}}-1$ and $T_{\mathrm{S}}=2^{R_{\mathrm{S}}}-1$. The following proposition presents the coverage probability of user U_0.

Proposition 1. *When both users U_0 and U_{p} are LOS, $\phi_{0\mathrm{p}}\leq\Theta_{\mathrm{B}}^{\mathrm{NOMA}}$, and $r_0\leq r_{\mathrm{p}}$, the coverage probability of typical user U_0 can be computed as*

$$P_{\mathrm{I},0}^{\mathrm{cov}}\left(\xi_1\right)=\sum_{k_{\mathrm{L}}=2}^{\infty}\sum_{k_{\mathrm{N}}=0}^{\infty}\frac{k_{\mathrm{L}}}{k_{\mathrm{L}}+k_{\mathrm{N}}}\Psi_{\mathrm{L}}(k_{\mathrm{L}})\Psi_{\mathrm{N}}(k_{\mathrm{N}})Q_{\mathrm{I},0}(\xi_1), \tag{26}$$

if $\alpha_{\mathrm{F}}^2>T_{\mathrm{F}}\alpha_{\mathrm{S}}^2$, otherwise $P_{\mathrm{I},0}^{\mathrm{cov}}\left(\xi_1\right)=0$, where $\xi_1=\max\left\{\frac{T_{\mathrm{F}}\sigma^2}{\left(\alpha_{\mathrm{F}}^2-T_{\mathrm{F}}\alpha_{\mathrm{S}}^2\right)P_{\mathrm{B}}},\frac{T_{\mathrm{S}}\sigma^2}{\alpha_{\mathrm{S}}^2P_{\mathrm{B}}}\right\}$, $Q_{\mathrm{I},0}(\xi_1)$ is given in (28), shown at the top of this page, $\phi_{0\mathrm{p},m_1}=\frac{\Theta_{\mathrm{B}}^{\mathrm{NOMA}}}{2}\left(\zeta_{m_1}+1\right)$, $\zeta_{m_1}=\cos\left(\frac{2m_1-1}{2M_1}\pi\right)$, $r_{\mathrm{n},m_2}=\frac{R}{2}\left(\varpi_{m_2}+1\right)$, $\varpi_{m_2}=\cos\left(\frac{2m_2-1}{2M_2}\pi\right)$, M_1 and M_2 are parameters for balancing the tradeoff between computational complexity and the accuracy of the Gauss-Chebyshev quadrature approximation, c_i and $d_i\left(\phi_{0\mathrm{p},m_1}\right)$ are given in (21) and (22), respectively, and $f_{r_{\mathrm{n}}}\left(r_{\mathrm{n},m_2}\right)$ and $f_{\phi_{0\mathrm{p}}}\left(\phi_{0\mathrm{p},m_1}\right)$ are given in (16) and (18), respectively.

Proof. Please refer to Appendix F. □

Note that, for the system model under consideration, the approximation in (28) is only due to the use of Gauss-Chebyshev quadrature, which approximates the integrals over the angle difference and the distance by summations. The total number of Gauss-Chebyshev quadrature summation terms used to approximate these two independent integrals is M_1+M_2. Numerical results (not shown here) reveal that the approximation error is less than 10^{-3} when $M_1=M_2=30$. According to Proposition 1, the coverage probability of typical user U_0 depends on the target data rates of both paired NOMA users. In addition, the coverage probability also depends on the beamwidth of the main lobe (i.e., $\Theta_{\mathrm{B}}^{\mathrm{NOMA}}$), as it determines the antenna gain of the main lobe and also affects the probability that typical user U_0 is covered by the main lobe of the base station in the presence of beam misalignment. The transmit power allocation coefficients should be appropriately set, i.e., $\alpha_{\mathrm{F}}^2>T_{\mathrm{F}}\alpha_{\mathrm{S}}^2$, so as to ensure that the typical user U_0 can successfully perform SIC, which is the prerequisite for the decoding of its own signal. Hence, the target data rate of the user whose signal is decoded first, R_{F}, is limited by the transmit power allocation coefficients, i.e., $R_{\mathrm{F}}<\log_2\left(1+\frac{\alpha_{\mathrm{F}}^2}{\alpha_{\mathrm{S}}^2}\right)$. On the other hand, a small value of α_{S}^2 limits the achievable rate of the user whose signal is decoded second, R_{S}. Hence, the value of $\alpha_{\mathrm{F}}^2\in\left(\frac{T_{\mathrm{F}}}{1+T_{\mathrm{F}}},1\right)$ can be optimized to enhance the outage sum rate.

The following corollary provides a simplified expression for $Q_{\mathrm{I},0}(\xi_1)$ in (27) for the case that the LOS ball model [18], [33] is used and beamsteering errors are absent. Note that the LOS probability of the LOS ball model is modeled as a step function, i.e., $p(r)=\mathbb{1}(r<R_{\mathrm{LOS}})$, where $\mathbb{1}(\cdot)$ denotes the indicator function and $R_{\mathrm{LOS}}<R$ is the maximum length of a LOS link.

Corollary 1. *If the LOS ball model is employed and beamsteering errors are absent, $Q_{\mathrm{I},0}(\xi_1)$ in (27) can be simplified as*

$$Q_{\mathrm{I},0}(\xi_1)=\left(1-\left(1-\frac{\Theta_{\mathrm{B}}^{\mathrm{NOMA}}}{2\pi}\right)^{2k_{\mathrm{L}}-2}\right)\sum_{n=0}^{N_{\mathrm{L}}-1}\frac{4}{n!\beta_{\mathrm{L}}R_{\mathrm{LOS}}^2}$$

$$\times\left(\left(\frac{N_{\mathrm{L}}\xi_1}{c_1C_{\mathrm{L}}}\right)^{\beta_{\mathrm{L}}}\gamma\left(n+\frac{1}{\beta_{\mathrm{L}}},\frac{N_{\mathrm{L}}\xi_1}{c_1C_{\mathrm{L}}}R_{\mathrm{LOS}}^{\beta_{\mathrm{L}}}\right)\right.$$

$$\left.-\frac{1}{R_{\mathrm{LOS}}^2}\left(\frac{N_{\mathrm{L}}\xi_1}{c_1C_{\mathrm{L}}}\right)^{\beta_{\mathrm{L}}/4}\gamma\left(n+\frac{4}{\beta_{\mathrm{L}}},\frac{N_{\mathrm{L}}\xi_1}{c_1C_{\mathrm{L}}}R_{\mathrm{LOS}}^{\beta_{\mathrm{L}}}\right)\right), \tag{29}$$

where $\gamma(\cdot,\cdot)$ is the lower incomplete Gamma function.

Proof. Please refer to Appendix G. □

Without beamsteering errors, $Q_{\mathrm{I},0}(\xi_1)$ increases linearly with the probability that the angle difference $\phi_{0\mathrm{p}}$ is smaller than $\Theta_{\mathrm{B}}^{\mathrm{NOMA}}$, i.e., with $1-\left(1-\frac{\Theta_{\mathrm{B}}^{\mathrm{NOMA}}}{2\pi}\right)^{2k_{\mathrm{L}}-2}$, as this is the probability that both NOMA users are covered by the main lobe of the base station. In NOMA-I, paired NOMA user U_{p} is the far LOS user. Thereby, user U_{p} can successfully decode its own signal if the SINR of its own signal is not smaller than T_{F} (i.e., $\gamma_{\mathrm{p}|0}\geq T_{\mathrm{F}}$). The following proposition provides the coverage probability for user U_{p}.

Proposition 2. *When both users U_0 and U_{p} are LOS, $\phi_{0\mathrm{p}}\leq\Theta_{\mathrm{B}}^{\mathrm{NOMA}}$, and $r_0\leq r_{\mathrm{p}}$, the coverage probability of user U_{p} is given by*

$$P_{\mathrm{I,p}}^{\mathrm{cov}}\left(\xi_2\right)=\sum_{k_{\mathrm{L}}=2}^{\infty}\sum_{k_{\mathrm{N}}=0}^{\infty}\frac{k_{\mathrm{L}}}{k_{\mathrm{L}}+k_{\mathrm{N}}}\Psi_{\mathrm{L}}(k_{\mathrm{L}})\Psi_{\mathrm{N}}(k_{\mathrm{N}})Q_{\mathrm{I,p}}(\xi_2), \tag{30}$$

if $\alpha_{\mathrm{F}}^2 > T_{\mathrm{F}}\alpha_{\mathrm{S}}^2$, *otherwise* $P_{\mathrm{I,p}}^{\mathrm{cov}}(\xi_2) = 0$, *where* $\xi_2 = \frac{T_{\mathrm{F}}\sigma^2}{(\alpha_{\mathrm{F}}^2 - T_{\mathrm{F}}\alpha_{\mathrm{S}}^2)P_{\mathrm{B}}}$, $Q_{\mathrm{I,p}}(\xi_2)$ *is given in (32), shown at the top of the next page,* $r_{\mathrm{f},m_2} = \frac{R}{2}(\varpi_{m_2}+1)$, ζ_{m_1}, $\phi_{0\mathrm{p},m_1}$, *and* ϖ_{m_2} *are given in Proposition 1,* $f_{r_{\mathrm{f}}}(r_{\mathrm{f},m_2})$ *is given in (17), and* c_i *and* $v_i(\phi_{0\mathrm{p},m_1})$ *are given in (21) and (23), respectively.*

Proof. Please refer to Appendix H. □

According to Proposition 2, the coverage probability of paired NOMA user U_{p} does not depend on the target data rate of typical user U_0 as user U_{p} does not need to perform SIC. The coverage probability does depend on the PDF of the distance of the far LOS user and the transmit power allocation coefficients.

2) NOMA-II: The case of NOMA-II occurs when both paired NOMA users are LOS, $\phi_{0\mathrm{p}} \le \Theta_{\mathrm{B}}^{\mathrm{NOMA}}$, and $r_0 > r_{\mathrm{p}}$. In this case, users U_0 and U_{p} are the far and near LOS users, respectively. Hence, the coverage probabilities of users U_0 and U_{p} can, respectively, be expressed as

$$P_{\mathrm{II},0}^{\mathrm{cov}}(\xi_2) = \sum_{k_{\mathrm{L}}=2}^{\infty}\sum_{k_{\mathrm{N}}=0}^{\infty}\frac{k_{\mathrm{L}}\Psi_{\mathrm{L}}(k_{\mathrm{L}})}{k_{\mathrm{L}}+k_{\mathrm{N}}}\Psi_{\mathrm{N}}(k_{\mathrm{N}})Q_{\mathrm{II},0}(\xi_2), \quad (33)$$

$$P_{\mathrm{II,p}}^{\mathrm{cov}}(\xi_1) = \sum_{k_{\mathrm{L}}=2}^{\infty}\sum_{k_{\mathrm{N}}=0}^{\infty}\frac{k_{\mathrm{L}}\Psi_{\mathrm{L}}(k_{\mathrm{L}})}{k_{\mathrm{L}}+k_{\mathrm{N}}}\Psi_{\mathrm{N}}(k_{\mathrm{N}})Q_{\mathrm{II,p}}(\xi_1), \quad (34)$$

where

$$Q_{\mathrm{II},0}(\xi_2) = \mathbb{P}\left(\gamma_{0|\mathrm{p}} \ge T_{\mathrm{F}}, \phi_{0\mathrm{p}} \le \Theta_{\mathrm{B}}^{\mathrm{NOMA}}, r_0 > r_{\mathrm{p}}\right), \quad (35)$$

$$Q_{\mathrm{II,p}}(\xi_1) = \mathbb{P}\left(\gamma_{0\to\mathrm{p}} \ge T_{\mathrm{F}}, \gamma_{\mathrm{p}} \ge T_{\mathrm{S}}, \phi_{0\mathrm{p}} \le \Theta_{\mathrm{B}}^{\mathrm{NOMA}}, r_{\mathrm{p}} < r_0\right). \quad (36)$$

Following the same steps as in the proofs of Propositions 1 and 2, we obtain $Q_{\mathrm{II},0}(\xi_2)$ from $Q_{\mathrm{I,p}}(\xi_2)$ given in (31) by replacing $v_i(\phi_{0\mathrm{p},m_1})$ with $d_i(\phi_{0\mathrm{p},m_1})$, and we obtain $Q_{\mathrm{II,p}}(\xi_1)$ from $Q_{\mathrm{I},0}(\xi_1)$ given in (27) by replacing $d_i(\phi_{0\mathrm{p},m_1})$ with $v_i(\phi_{0\mathrm{p},m_1})$. Note that, in (33) and (34), $\sum_{k_{\mathrm{L}}=2}^{\infty}\sum_{k_{\mathrm{N}}=0}^{\infty}\frac{k_{\mathrm{L}}}{k_{\mathrm{L}}+k_{\mathrm{N}}}\Psi_{\mathrm{L}}(k_{\mathrm{L}})\Psi_{\mathrm{N}}(k_{\mathrm{N}})$ is the probability that there are at least two LOS users and typical user U_0 is LOS, and $Q_{\mathrm{II,p}}(\xi_1)$ is the probability that paired NOMA user U_{p} can successfully perform SIC and decode its own signal when NOMA is enabled.

3) NOMA-III: The case of NOMA-III occurs when typical user U_0 is NLOS and paired NOMA user U_{p} is LOS, and $\phi_{0\mathrm{p}} \le \Theta_{\mathrm{B}}^{\mathrm{NOMA}}$. In this case, the signal intended for user U_0 is decoded first. The coverage probabilities of users U_0 and U_{p} can, respectively, be expressed as

$$P_{\mathrm{III},0}^{\mathrm{cov}}(\xi_2) = \sum_{k_{\mathrm{L}}=1}^{\infty}\sum_{k_{\mathrm{N}}=1}^{\infty}\frac{k_{\mathrm{N}}\Psi_{\mathrm{L}}(k_{\mathrm{L}})}{k_{\mathrm{L}}+k_{\mathrm{N}}}\Psi_{\mathrm{N}}(k_{\mathrm{N}})Q_{\mathrm{III},0}(\xi_2), \quad (37)$$

$$P_{\mathrm{III,p}}^{\mathrm{cov}}(\xi_1) = \sum_{k_{\mathrm{L}}=1}^{\infty}\sum_{k_{\mathrm{N}}=1}^{\infty}\frac{k_{\mathrm{N}}\Psi_{\mathrm{L}}(k_{\mathrm{L}})}{k_{\mathrm{L}}+k_{\mathrm{N}}}\Psi_{\mathrm{N}}(k_{\mathrm{N}})Q_{\mathrm{III,p}}(\xi_1), \quad (38)$$

where

$$Q_{\mathrm{III},0}(\xi_2) = \mathbb{P}\left(\gamma_{0|\mathrm{p}} \ge T_{\mathrm{F}}, \phi_{0\mathrm{p}} \le \Theta_{\mathrm{B}}^{\mathrm{NOMA}}\right), \quad (39)$$

$$Q_{\mathrm{III,p}}(\xi_1) = \mathbb{P}\left(\gamma_{0\to\mathrm{p}} \ge T_{\mathrm{F}}, \gamma_{\mathrm{p}} \ge T_{\mathrm{S}}, \phi_{0\mathrm{p}} \le \Theta_{\mathrm{B}}^{\mathrm{NOMA}}\right). \quad (40)$$

As typical user U_0 is NLOS and is randomly selected, we obtain $Q_{\mathrm{III},0}(\xi_2)$ from $Q_{\mathrm{I},0}(\xi_1)$ given in (27) by replacing N_{L}, ξ_1, C_{L}, β_{L}, and $f_{r_{\mathrm{n}}}(r_{\mathrm{n},m_2})$ by N_{N}, ξ_2, C_{N}, β_{N}, and $f_{\mathrm{N}}(r_{\mathrm{n},m_2})$, respectively. Similarly, we obtain $Q_{\mathrm{III,p}}(\xi_1)$ from $Q_{\mathrm{I,p}}(\xi_2)$ given in (31) by replacing ξ_2 and $f_{r_{\mathrm{f}}}(r_{\mathrm{f},m_2})$ with ξ_1 and $f_{\mathrm{L}}(r_{\mathrm{f},m_2})$, respectively. Note that $\sum_{k_{\mathrm{L}}=1}^{\infty}\sum_{k_{\mathrm{N}}=1}^{\infty}\frac{k_{\mathrm{N}}}{k_{\mathrm{L}}+k_{\mathrm{N}}}\Psi_{\mathrm{L}}(k_{\mathrm{L}})\Psi_{\mathrm{N}}(k_{\mathrm{N}})$ is the probability that there is at least one LOS user and typical user U_0 is a NLOS user.

C. Coverage Probability of OMA Transmission

OMA transmission is enabled when a second NOMA user for pairing does not exist or the angle difference is $\phi_{0\mathrm{p}} > \Theta_{\mathrm{B}}^{\mathrm{NOMA}}$. Based on Lemma 5, the complementary CDF (CCDF) of SNR $\gamma_{j,\nu}^{\mathrm{OMA}}, j \in \{0,\mathrm{p}\}, \nu \in \{\mathrm{L},\mathrm{N}\}$, defined in (12) is given by

$$\begin{aligned}
\overline{F}_{\gamma_{j,\nu}^{\mathrm{OMA}}}(\gamma) &\overset{(a)}{=} \mathbb{P}\left(|h_j|^2 P_{\mathrm{B}}\mathcal{D}_j^{\mathrm{OMA}}\ell_\nu(r_j)/\sigma^2 \ge \gamma\right)\\
&\overset{(b)}{=} \sum_{n=0}^{N_\nu-1}\frac{\left(N_\nu\gamma\sigma^2/P_{\mathrm{B}}\right)^n}{n!}\mathbb{E}_{\{\mathcal{D}_j^{\mathrm{OMA}},r_j\}}\left[\frac{\exp\left(-\frac{N_\nu\gamma\sigma^2/P_{\mathrm{B}}}{\mathcal{D}_j^{\mathrm{OMA}}\ell_\nu(r_j)}\right)}{\left(\mathcal{D}_j^{\mathrm{OMA}}\ell_\nu(r_j)\right)^n}\right]\\
&\overset{(c)}{=} \sum_{n=0}^{N_\nu-1}\frac{\left(N_\nu\gamma\sigma^2/P_{\mathrm{B}}\right)^n}{n!}\sum_{i=1}^{4}\tau_i\int_0^R\left(\frac{r_j^{\beta_\nu}}{l_iC_\nu}\right)^n\\
&\quad\times\exp\left(-\frac{N_\nu\gamma\sigma^2 r_j^{\beta_\nu}}{P_{\mathrm{B}}l_iC_\nu}\right)f_\nu(r_j)\mathrm{d}r_j,\\
&\overset{(d)}{\approx} \sum_{n=0}^{N_\nu-1}\frac{R\pi\left(N_\nu\gamma\sigma^2/P_{\mathrm{B}}\right)^n}{2M_2 n!}\sum_{i=1}^{4}\tau_i\sum_{m_2=1}^{M_2}\sqrt{1-\varpi_{m_2}^2}\\
&\quad\times\left(\frac{r_{j,m_2}^{\beta_\nu}}{l_iC_\nu}\right)^n\exp\left(-\frac{N_\nu\gamma\sigma^2 r_{j,m_2}^{\beta_\nu}}{P_{\mathrm{B}}l_iC_\nu}\right)f_\nu(r_{j,m_2}),
\end{aligned} \quad (41)$$

where (a) follows by substituting (12), (b) follows from the normalized Gamma distribution of $|h_j|^2$, (c) follows by taking the expectations over $\mathcal{D}_j^{\mathrm{OMA}}$ and r_j, and (d) follows by applying Gauss-Chebyshev quadrature [38]. Note that we adopt an exact expression for the CDF of the normalized random variable $|h_j|^2$, rather than its upper bound as in previous works [18].

When typical user U_0 is LOS, its coverage probability can be computed as

$$\begin{aligned}
P_{\mathrm{OMA},0}^{\mathrm{cov1}}(T_{\mathrm{F}}) &= \Psi_{\mathrm{L}}(1)\Psi_{\mathrm{N}}(0)\overline{F}_{\gamma_{0,\mathrm{L}}^{\mathrm{OMA}}}(T_{\mathrm{F}})\\
&\quad+\Psi_{\mathrm{L}}(1)\sum_{k_{\mathrm{N}}=1}^{\infty}\frac{\Psi_{\mathrm{N}}(k_{\mathrm{N}})}{1+k_{\mathrm{N}}}\overline{F}_{\gamma_{0,\mathrm{L}}^{\mathrm{OMA}}}(2T_{\mathrm{F}})\\
&\quad+\sum_{k_{\mathrm{L}}=2}^{\infty}\sum_{k_{\mathrm{N}}=0}^{\infty}\frac{k_{\mathrm{L}}}{k_{\mathrm{L}}+k_{\mathrm{N}}}\Psi_{\mathrm{L}}(k_{\mathrm{L}})\Psi_{\mathrm{N}}(k_{\mathrm{N}})Q_1^{\mathrm{OMA}}(T_{\mathrm{F}}),
\end{aligned} \quad (42)$$

where $Q_1^{\mathrm{OMA}}(T_{\mathrm{F}}) = \mathbb{P}\left(\gamma_{0,\mathrm{L}}^{\mathrm{OMA}} \ge 2T_{\mathrm{F}}, \phi_{0\mathrm{p}} > \Theta_{\mathrm{B}}^{\mathrm{NOMA}}\right) \overset{(a)}{=} \overline{F}_{\gamma_{0,\mathrm{L}}^{\mathrm{OMA}}}(2T_{\mathrm{F}})\left(\frac{2\pi-\Theta_{\mathrm{B}}^{\mathrm{NOMA}}}{2\pi}\right)^{2k_{\mathrm{L}}-2}$. Note that (a) follows from the independence between SNR $\gamma_{0,\mathrm{L}}^{\mathrm{OMA}}$ and angle difference $\phi_{0\mathrm{p}}$. On the right hand side of (42), the first and second terms represent the coverage probability when user U_0 is the only LOS user in the network, where a second NOMA user for

$$Q_{\mathrm{I,p}}(\xi_2)=\sum_{n=0}^{N_{\mathrm{L}}-1}\sum_{i=1}^{4}\frac{(N_{\mathrm{L}}\xi_2)^n}{n!\,(c_iC_{\mathrm{L}})^n}\int_0^{\Theta_{\mathrm{B}}^{\mathrm{NOMA}}}v_i(\phi)f_{\phi_{0\mathrm{p}}}(\phi)\mathrm{d}\phi\int_0^R r^{n\beta_{\mathrm{L}}}\exp\left(-\frac{N_{\mathrm{L}}\xi_2r^{\beta_{\mathrm{L}}}}{c_iC_{\mathrm{L}}}\right)f_{r_{\mathrm{f}}}(r)\mathrm{d}r \tag{31}$$

$$\approx\sum_{n=0}^{N_{\mathrm{L}}-1}\sum_{i=1}^{4}\frac{\Theta_{\mathrm{B}}^{\mathrm{NOMA}}R\pi^2\,(N_{\mathrm{L}}\xi_2)^n}{4M_1M_2n!}\left(\sum_{m_1=1}^{M_1}\sqrt{1-\zeta_{m_1}^2}v_i\left(\phi_{0\mathrm{p},m_1}\right)f_{\phi_{0\mathrm{p}}}\left(\phi_{0\mathrm{p},m_1}\right)\right)$$
$$\times\sum_{m_2=1}^{M_2}\sqrt{1-\varpi_{m_2}^2}\left(\frac{r_{\mathrm{f},m_2}^{\beta_{\mathrm{L}}}}{c_iC_{\mathrm{L}}}\right)^n\exp\left(-\frac{N_{\mathrm{L}}\xi_2r_{\mathrm{f},m_2}^{\beta_{\mathrm{L}}}}{c_iC_{\mathrm{L}}}\right)f_{r_{\mathrm{f}}}\left(r_{\mathrm{f},m_2}\right), \tag{32}$$

pairing does not exist. The third term is the coverage probability when both users U_0 and U_{p} are LOS but $\phi_{0\mathrm{p}}>\Theta_{\mathrm{B}}^{\mathrm{NOMA}}$. When OMA is enabled, the base station transmits the signals to users U_0 and U_{p} in the first and second halves of a time slot, respectively. Note that $Q_1^{\mathrm{OMA}}(T_{\mathrm{F}})$ is a decreasing function of the beamwidth of the main lobe of the base station ($\Theta_{\mathrm{B}}^{\mathrm{NOMA}}$) and the number of LOS users (k_{L}).

On the other hand, when typical user U_0 is NLOS, its coverage probability is given by

$$\begin{aligned}P_{\mathrm{OMA},0}^{\mathrm{cov2}}(T_{\mathrm{F}})&=\Psi_{\mathrm{L}}(0)\Psi_{\mathrm{N}}(1)\overline{F}_{\gamma_{0,\mathrm{N}}^{\mathrm{OMA}}}(T_{\mathrm{F}})\\&+\Psi_{\mathrm{L}}(0)\left(1-\Psi_{\mathrm{N}}(0)-\Psi_{\mathrm{N}}(1)\right)\overline{F}_{\gamma_{0,\mathrm{N}}^{\mathrm{OMA}}}(2T_{\mathrm{F}})\\&+\sum_{k_{\mathrm{L}}=1}^{\infty}\sum_{k_{\mathrm{N}}=1}^{\infty}\frac{k_{\mathrm{N}}\Psi_{\mathrm{L}}(k_{\mathrm{L}})}{k_{\mathrm{L}}+k_{\mathrm{N}}}\Psi_{\mathrm{N}}(k_{\mathrm{N}})Q_2^{\mathrm{OMA}}(T_{\mathrm{F}}),\end{aligned} \tag{43}$$

where $Q_2^{\mathrm{OMA}}(T_{\mathrm{F}})=\mathbb{P}\left(\gamma_{0,\mathrm{N}}^{\mathrm{OMA}}\geq 2T_{\mathrm{F}},\phi_{0\mathrm{p}}>\Theta_{\mathrm{B}}^{\mathrm{NOMA}}\right)=\overline{F}_{\gamma_{0,\mathrm{N}}^{\mathrm{OMA}}}(2T_{\mathrm{F}})\left(\frac{2\pi-\Theta_{\mathrm{B}}^{\mathrm{NOMA}}}{2\pi}\right)^{2k_{\mathrm{L}}}$.

On the right hand side of (43), the first and second terms represent the coverage probability when there are no LOS users in the network coverage area. The third term is the coverage probability when user U_0 is NLOS and user U_{p} is LOS but $\phi_{0\mathrm{p}}>\Theta_{\mathrm{B}}^{\mathrm{NOMA}}$.

Similar to (42) and (43), when OMA is used, the coverage probability of user U_{p} is given by

$$\begin{aligned}P_{\mathrm{OMA,p}}^{\mathrm{cov}}(T_{\mathrm{S}})&=\Psi_{\mathrm{L}}(1)\sum_{k_{\mathrm{N}}=1}^{\infty}\frac{\Psi_{\mathrm{N}}(k_{\mathrm{N}})}{1+k_{\mathrm{N}}}\overline{F}_{\gamma_{\mathrm{p,N}}^{\mathrm{OMA}}}(2T_{\mathrm{S}})\\&+\sum_{k_{\mathrm{L}}=2}^{\infty}\sum_{k_{\mathrm{N}}=0}^{\infty}\frac{k_{\mathrm{L}}\Psi_{\mathrm{L}}(k_{\mathrm{L}})}{k_{\mathrm{L}}+k_{\mathrm{N}}}\Psi_{\mathrm{N}}(k_{\mathrm{N}})Q_1^{\mathrm{OMA}}(T_{\mathrm{S}})\\&+\Psi_{\mathrm{L}}(0)\left(1-\Psi_{\mathrm{N}}(0)-\Psi_{\mathrm{N}}(1)\right)\overline{F}_{\gamma_{\mathrm{p,N}}^{\mathrm{OMA}}}(2T_{\mathrm{S}})\\&+\sum_{k_{\mathrm{L}}=1}^{\infty}\sum_{k_{\mathrm{N}}=1}^{\infty}\frac{k_{\mathrm{N}}\Psi_{\mathrm{L}}(k_{\mathrm{L}})}{k_{\mathrm{L}}+k_{\mathrm{N}}}\Psi_{\mathrm{N}}(k_{\mathrm{N}})Q_3^{\mathrm{OMA}}(T_{\mathrm{S}}),\end{aligned} \tag{44}$$

where $Q_3^{\mathrm{OMA}}(T_{\mathrm{S}})=\overline{F}_{\gamma_{\mathrm{p,L}}^{\mathrm{OMA}}}(2T_{\mathrm{S}})\left(\frac{2\pi-\Theta_{\mathrm{B}}^{\mathrm{NOMA}}}{2\pi}\right)^{2k_{\mathrm{L}}}$.

According to (42), (43), and (44), the probability that OMA transmission is enabled depends on the user density (λ), the average LOS range (η), and the beamwidth of the main lobe ($\Theta_{\mathrm{B}}^{\mathrm{NOMA}}$).

D. Coverage Probability and Outage Sum Rate of Proposed NOMA Scheme

Based on the analysis in Sections III-B and III-C, we present the main result on the coverage probability of the proposed NOMA scheme in the following theorem.

Theorem 1. *The coverage probabilities of users U_0 and U_{p} for the proposed NOMA scheme in mmWave networks with beam misalignment can, respectively, be expressed as*

$$\begin{aligned}P_0^{\mathrm{cov}}&=0.5\left(P_{\mathrm{I},0}^{\mathrm{cov}}(\xi_1)+P_{\mathrm{II},0}^{\mathrm{cov}}(\xi_2)\right)+P_{\mathrm{III},0}^{\mathrm{cov}}(\xi_2)\\&+P_{\mathrm{OMA},0}^{\mathrm{cov1}}(T_{\mathrm{F}})+P_{\mathrm{OMA},0}^{\mathrm{cov2}}(T_{\mathrm{F}}),\end{aligned} \tag{45}$$

$$\begin{aligned}P_{\mathrm{p}}^{\mathrm{cov}}&=0.5\left(P_{\mathrm{I,p}}^{\mathrm{cov}}(\xi_2)+P_{\mathrm{II,p}}^{\mathrm{cov}}(\xi_1)\right)+P_{\mathrm{III,p}}^{\mathrm{cov}}(\xi_1)\\&+P_{\mathrm{OMA,p}}^{\mathrm{cov}}(T_{\mathrm{S}}).\end{aligned} \tag{46}$$

Proof. When both users U_0 and U_{p} are LOS, they have the same probability (i.e., 0.5) to be the near or the far LOS user due to the randomness of the user locations. By considering all the cases for NOMA and OMA transmission in Sections III-B and III-C, respectively, we directly obtain the coverage probabilities of users U_0 and U_{p} given in (45) and (46), respectively. $\square$

Outage rate is defined as the mean data rate achievable at a user when its signal is transmitted with a certain target data rate. Based on the derived coverage probability, the outage sum rate of the typical user and the paired NOMA user can be directly calculated as

$$\begin{aligned}R_{\mathrm{sum}}=&\left(\frac{1}{2}\left(P_{\mathrm{I,p}}^{\mathrm{cov}}(\xi_1)+P_{\mathrm{II},0}^{\mathrm{cov}}(\xi_2)\right)+P_{\mathrm{III},0}^{\mathrm{cov}}(\xi_2)\right)R_{\mathrm{F}}\\&+\left(P_{\mathrm{OMA},0}^{\mathrm{cov1}}(T_{\mathrm{F}})+P_{\mathrm{OMA},0}^{\mathrm{cov2}}(T_{\mathrm{F}})\right)R_{\mathrm{F}}\\&+\left(\frac{1}{2}\left(P_{\mathrm{I},0}^{\mathrm{cov}}(\xi_2)+P_{\mathrm{II,p}}^{\mathrm{cov}}(\xi_1)\right)+P_{\mathrm{III,p}}^{\mathrm{cov}}(\xi_1)\right)R_{\mathrm{S}}\\&+P_{\mathrm{OMA,p}}^{\mathrm{cov}}(T_{\mathrm{S}})R_{\mathrm{S}}.\end{aligned} \tag{47}$$

The outage sum rate is not a monotonic function of the radius of the network coverage area, R. As the polynomially increasing term with respect to R in the expression for the coverage probability (e.g., (28)) cannot compensate for the exponentially decreasing term when R is large, the outage sum rate first increases and then decreases as R increases. Hence, the value of R can be optimized to enhance the spectrum efficiency.

IV. Ergodic Sum Rate Analysis

In this section, we derive the ergodic sum rate of both paired NOMA users for the proposed scheme in mmWave networks. The ergodic rate, as an important metric of spectral efficiency, refers to the mean data rate achievable at a user when adaptive modulation/coding is used to achieve the Shannon bound for a given instantaneous SINR [39]. The ergodic rate is given by

$$\mathcal{R}=\mathbb{E}_{\gamma}\left[\log_2\left(1+\min\{\gamma,T_{\max}\}\right)\right]$$

$$= \frac{1}{\ln 2} \int_0^{T_{\max}} \frac{\overline{F}_\gamma(x)}{1+x} \mathrm{d}x, \tag{48}$$

where γ and $\overline{F}_\gamma(x)$ denote the SINR and its CCDF, respectively, $\mathbb{E}_\gamma[\cdot]$ is the expectation with respect to SINR γ, $T_{\max} = 2^{R_{\max}} - 1$ is the SINR threshold imposed by practical constraints for the RF circuit, and $R_{\max}$ is the maximum achievable rate. If the SINR exceeds $T_{\max}$, this cannot be exploited for further rate improvement. Instead of using fixed target data rates as in the analysis of the coverage probability, the ergodic rate is opportunistically determined by the instantaneous SINR observed at the user and SINR threshold $T_{\max}$. In this paper, the ergodic sum rate refers to the summation of the ergodic rates of typical user U_0 and paired NOMA user U_p. In the following subsections, we analyze the ergodic sum rates for NOMA and OMA transmission.

A. Ergodic Sum Rate of NOMA Transmission

As demonstrated in [40], when the far user can successfully decode its own signal, the near user can always decode the signal intended for the far user due to its better channel conditions. Hence, the ergodic rates of users U_0 and U_p in NOMA-I are determined by the SINRs γ_0 and $\gamma_{\mathrm{p}|0}$ given in (6) and (7), respectively. The following proposition presents the ergodic sum rate of the paired NOMA users in NOMA-I.

Proposition 3. *When both users U_0 and U_p are LOS, $\phi_{0\mathrm{p}} \leq \Theta_\mathrm{B}^{\mathrm{NOMA}}$, and $r_0 \leq r_\mathrm{p}$, the ergodic sum rate of the proposed NOMA scheme can be computed as*

$$\mathcal{R}_\mathrm{I}^{\mathrm{NOMA}} \approx \frac{T_{\max}\pi}{2M_3 \ln 2} \sum_{m_3=1}^{M_3} \sqrt{1-\delta_{m_3}^2} \frac{P_{\mathrm{I},0}^{\mathrm{cov}}(\xi_3(\gamma_{m_3}))}{1+\gamma_{m_3}} + \frac{\min\left\{T_{\max}, \frac{\alpha_\mathrm{F}^2}{\alpha_\mathrm{S}^2}\right\}\pi}{2M_4 \ln 2} \sum_{m_4=1}^{M_4} \sqrt{1-\chi_{m_4}^2} \frac{P_{\mathrm{I},\mathrm{p}}^{\mathrm{cov}}(\xi_4(\gamma_{m_4}))}{1+\gamma_{m_4}}, \tag{49}$$

where $\xi_3(\gamma_{m_3}) = \frac{\sigma^2\gamma_{m_3}}{\alpha_\mathrm{S}^2 P_\mathrm{B}}$, $\xi_4(\gamma_{m_4}) = \frac{\sigma^2\gamma_{m_4}}{(\alpha_\mathrm{F}^2 - \gamma_{m_4}\alpha_\mathrm{S}^2)P_\mathrm{B}}$, $\gamma_{m_3} = \frac{T_{\max}}{2}(\delta_{m_3}+1)$, $\delta_{m_3} = \cos\left(\frac{2m_3-1}{2M_3}\pi\right)$, $\gamma_{m_4} = \min\left\{T_{\max}, \frac{\alpha_\mathrm{F}^2}{\alpha_\mathrm{S}^2}\right\}(\chi_{m_4}+1)/2$, $\chi_{m_4} = \cos\left(\frac{2m_4-1}{2M_4}\pi\right)$, *$M_3$ and M_4 are parameters to balance the tradeoff between computational complexity and the accuracy of the Gauss-Chebyshev quadrature approximation, and $P_{\mathrm{I},0}^{\mathrm{cov}}(\xi_3(\gamma_{m_3}))$ and $P_{\mathrm{I},\mathrm{p}}^{\mathrm{cov}}(\xi_4(\gamma_{m_4}))$ are given by (26) and (30).*

Proof. Please refer to Appendix I. □

According to Proposition 3, the ergodic rate of each user is expressed as a function of the coverage probability. In addition, the ergodic rate of the far LOS user is limited by both the SINR threshold $T_{\max}$ and $\frac{\alpha_\mathrm{F}^2}{\alpha_\mathrm{S}^2}$, which indicates the importance of appropriate transmit power allocation. On the other hand, the ergodic rate of the near user is not limited by $\frac{\alpha_\mathrm{F}^2}{\alpha_\mathrm{S}^2}$. Thereby, the near user can achieve a higher ergodic rate. The following corollary provides the ergodic sum rate of the paired NOMA users in NOMA-II and NOMA-III.

Corollary 2. *When both users U_0 and U_p are LOS, $\phi_{0\mathrm{p}} \leq \Theta_\mathrm{B}^{\mathrm{NOMA}}$, and $r_0 > r_\mathrm{p}$, the ergodic sum rate of the proposed NOMA scheme is given by*

$$\mathcal{R}_{\mathrm{II}}^{\mathrm{NOMA}} \approx \frac{\min\left\{T_{\max}, \frac{\alpha_\mathrm{F}^2}{\alpha_\mathrm{S}^2}\right\}\pi}{2M_4 \ln 2} \sum_{m_4=1}^{M_4} \sqrt{1-\chi_{m_4}^2} \frac{P_{\mathrm{II},0}^{\mathrm{cov}}(\xi_4(\gamma_{m_4}))}{1+\gamma_{m_4}} + \frac{T_{\max}\pi}{2M_3 \ln 2} \sum_{m_3=1}^{M_3} \sqrt{1-\delta_{m_3}^2} \frac{P_{\mathrm{II},\mathrm{p}}^{\mathrm{cov}}(\xi_3(\gamma_{m_3}))}{1+\gamma_{m_3}}, \tag{50}$$

where $P_{\mathrm{II},0}^{\mathrm{cov}}(\xi_4(\gamma_{m_4}))$ and $P_{\mathrm{II},\mathrm{p}}^{\mathrm{cov}}(\xi_3(\gamma_{m_3}))$ are given by (33) and (34), respectively.

Similarly, when typical user U_0 is NLOS, paired NOMA user U_p is LOS, and $\phi_{0\mathrm{p}} \leq \Theta_\mathrm{B}^{\mathrm{NOMA}}$, the ergodic sum rate of the proposed NOMA scheme is given by

$$\mathcal{R}_{\mathrm{III}}^{\mathrm{NOMA}} \approx \frac{\min\left\{T_{\max}, \frac{\alpha_\mathrm{F}^2}{\alpha_\mathrm{S}^2}\right\}\pi}{2M_4 \ln 2} \sum_{m_4=1}^{M_4} \sqrt{1-\chi_{m_4}^2} \frac{P_{\mathrm{III},0}^{\mathrm{cov}}(\xi_4(\gamma_{m_4}))}{1+\gamma_{m_4}} + \frac{T_{\max}\pi}{2M_3 \ln 2} \sum_{m_3=1}^{M_3} \sqrt{1-\delta_{m_3}^2} \frac{P_{\mathrm{III},\mathrm{p}}^{\mathrm{cov}}(\xi_3(\gamma_{m_3}))}{1+\gamma_{m_3}}, \tag{51}$$

where $P_{\mathrm{III},0}^{\mathrm{cov}}(\xi_4(\gamma_{m_4}))$ and $P_{\mathrm{III},\mathrm{p}}^{\mathrm{cov}}(\xi_3(\gamma_{m_3}))$ are given by (37) and (38), respectively.

Proof. For these two cases, the ergodic rates of users U_0 and U_p are determined by the SINRs $\gamma_{0|\mathrm{p}}$ and γ_p given in (11) and (10), respectively. By following the same steps as in the proof of Proposition 1, we obtain (50) and (51). □

B. Ergodic Sum Rate of OMA Transmission

In the proposed scheme, OMA is enabled when a second NOMA user for pairing does not exist or the angle difference is $\phi_{0\mathrm{p}} > \Theta_\mathrm{B}^{\mathrm{NOMA}}$. In this case, the ergodic rate of user $U_j, j \in \{0, \mathrm{p}\}$, is determined by SNR $\gamma_{j,\nu}^{\mathrm{OMA}}, \nu \in \{\mathrm{L}, \mathrm{N}\}$, given in (12). Based on the CCDF of SNR $\gamma_{j,\nu}^{\mathrm{OMA}}$ derived in (41), the ergodic sum rate of users U_0 and U_p for OMA transmission is given in (52), shown at the top of the next page, where $P_{\mathrm{OMA},0}^{\mathrm{cov1}}(\gamma_{m_3})$, $P_{\mathrm{OMA},0}^{\mathrm{cov2}}(\gamma_{m_3})$, and $P_{\mathrm{OMA},\mathrm{p}}^{\mathrm{cov}}(\gamma_{m_3})$ are given in (42), (43), and (44), respectively.

C. Ergodic Sum Rate of Proposed Scheme

Based on the analysis in Sections IV-A and IV-B, we present the main result on the ergodic sum rate of the proposed scheme in the following theorem.

Theorem 2. *The ergodic sum rate of typical user U_0 and paired NOMA user U_p for the proposed NOMA scheme in mmWave networks with beam misalignment can be computed as*

$$\mathcal{R}^{\mathrm{NOMA}} = \frac{1}{2}\left(\mathcal{R}_\mathrm{I}^{\mathrm{NOMA}} + \mathcal{R}_{\mathrm{II}}^{\mathrm{NOMA}}\right) + \mathcal{R}_{\mathrm{III}}^{\mathrm{NOMA}} + \mathcal{R}^{\mathrm{OMA}}, \tag{53}$$

where $\mathcal{R}_\mathrm{I}^{\mathrm{NOMA}}$, $\mathcal{R}_{\mathrm{II}}^{\mathrm{NOMA}}$, $\mathcal{R}_{\mathrm{III}}^{\mathrm{NOMA}}$, and $\mathcal{R}^{\mathrm{OMA}}$ are given in (49), (50), (51), and (52), respectively.

$$\mathcal{R}^{\mathrm{OMA}} = \frac{1}{\ln 2}\int_0^{T_{\max}} \frac{P_{\mathrm{OMA},0}^{\mathrm{cov1}}(x) + P_{\mathrm{OMA},0}^{\mathrm{cov2}}(x) + P_{\mathrm{OMA,p}}^{\mathrm{cov}}(x)}{1+x}\mathrm{d}x$$
$$\approx \frac{T_{\max}\pi}{2M_3 \ln 2}\sum_{m_3=1}^{M_3}\sqrt{1-\delta_{m_3}^2}\frac{P_{\mathrm{OMA},0}^{\mathrm{cov1}}(\gamma_{m_3}) + P_{\mathrm{OMA},0}^{\mathrm{cov2}}(\gamma_{m_3}) + P_{\mathrm{OMA,p}}^{\mathrm{cov}}(\gamma_{m_3})}{1+\gamma_{m_3}}, \tag{52}$$

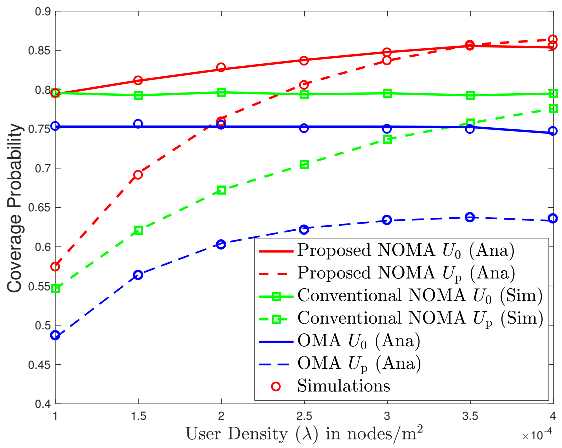


Fig. 3: Coverage probability of users U_0 and U_p versus user density.

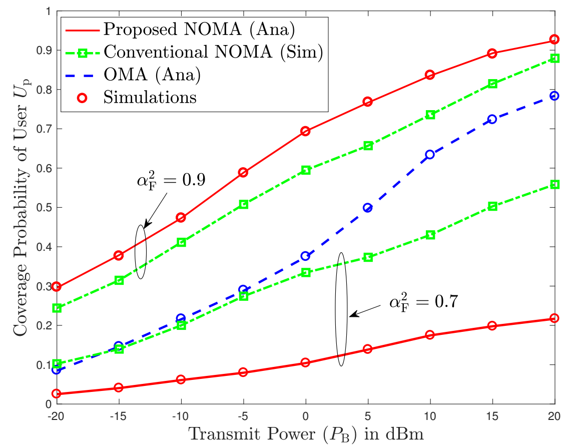


Fig. 4: Coverage probability of user U_p versus transmit power for $\lambda = 0.0003$ nodes/m^2.

Proof. When both users U_0 and U_p are LOS, they have the same probability (i.e., 0.5) to be the near and the far user. By utilizing the ergodic sum rates for NOMA and OMA transmission derived in (49), (50), (51), and (52), respectively, we obtain the ergodic sum rate of users U_0 and U_p given in (53). □

V. Performance Evaluation

In this section, we present the simulation and analytical results for the proposed NOMA scheme and compare them with the results for conventional NOMA and OMA. Conventional NOMA adopts distance-based user pairing, where the base station pairs the typical user with the user that is closest to the base station [37]. In OMA, for a fair comparison, the base station transmits the signals to the typical user and the paired NOMA user in the first and second halves of a time slot, respectively. In the simulations, we consider a circular network coverage area with radius $R = 200$ m. The noise power is -90 dBm. Unless specified otherwise, we set $\beta_\mathrm{L} = 2.5$, $\beta_\mathrm{N} = 3.5$, $C_\mathrm{L} = 2$, $C_\mathrm{N} = 1$, $N_\mathrm{L} = 3$, $N_\mathrm{N} = 2$, $\eta = 50$ m, $\alpha_\mathrm{F}^2 = 0.9$, $\alpha_\mathrm{S}^2 = 0.1$, $\varsigma = 0.1$, $\Theta_\mathrm{B}^{\mathrm{NOMA}} = \pi/3$, $\Theta_\mathrm{B}^{\mathrm{OMA}} = \pi/6$, $\Theta_\mathrm{U} = \pi/6$, $R_\mathrm{F} = 2$ bit per channel use (BPCU), $R_\mathrm{S} = 8$ BPCU, $P_\mathrm{B} = 10$ dBm, and $\sigma_\mathrm{B}^2 = \sigma_\mathrm{U}^2 = 0.2$. For all numerical results shown, we set the number of summation terms for the Gauss-Chebyshev quadrature to 50, which is sufficiently large to make the approximation error negligible.

Fig. 3 shows the impact of the user density (i.e., λ) on the coverage probability. The simulation (Sim) results match well with the analytical (Ana) results, which validates the performance analysis. As λ increases, the probability that the angle difference $\phi_{0\mathrm{p}}$ is smaller than $\Theta_\mathrm{B}^{\mathrm{NOMA}}$ increases, which in turn increases the probability that NOMA can be enabled and reduces the negative effects of beam misalignment. Hence, the coverage probabilities of typical user U_0 in the proposed NOMA scheme and paired NOMA user U_p in all considered schemes increase with λ. However, the coverage probability of typical user U_0 in conventional NOMA and OMA does not depend on λ, as the probability that typical user U_0 is served with NOMA is independent of λ. As the probability that both NOMA users are covered by the main lobe of the base station is higher, the coverage probabilities of both users are higher for the proposed NOMA scheme with angle-based user pairing compared to conventional NOMA with distance-based user pairing.

Fig. 4 illustrates the impact of the transmit power (i.e., P_B) and the power allocation coefficient (i.e., α_F^2) on the coverage probability of paired NOMA user U_p. As P_B increases, for all considered schemes the coverage probability of user U_p increases, as the received signal power also increases. When $\alpha_\mathrm{F}^2 = 0.9$, the proposed NOMA scheme outperforms conventional NOMA and OMA. However, if α_F^2 is decreased from 0.9 to 0.7, the coverage probabilities of user U_p under both NOMA schemes decrease significantly, and are even lower than the coverage probability of OMA. This is because the probability that paired NOMA user U_p can successfully perform SIC decreases with α_F^2. As the probability that OMA is enabled is higher, conventional NOMA with distance-based user pairing achieves a higher coverage probability than the proposed NOMA scheme with angle-based user pairing when the transmit power is not appropriately allocated. These results show that appropriate transmit power allocation plays an important role in exploiting the potential performance gains

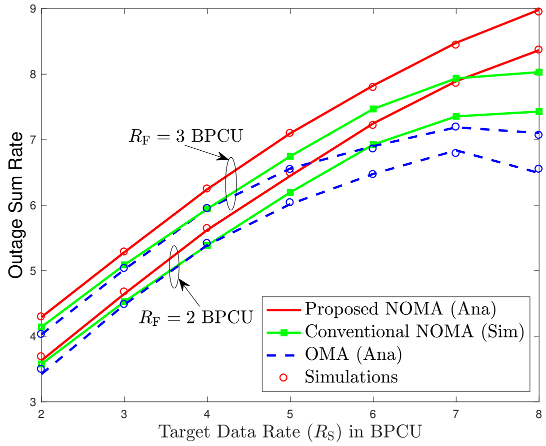


Fig. 5: Outage sum rate versus target data rates for $\lambda = 0.0003$ nodes/m^2.

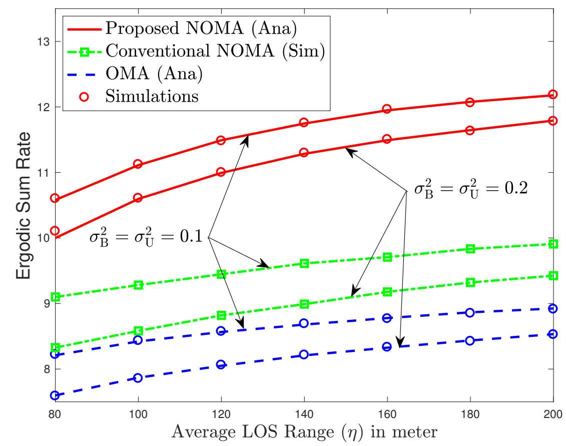


Fig. 7: Ergodic sum rate versus average LOS range for different beamsteering error variances, $\lambda = 0.0001$ nodes/m^2, $\beta_\mathrm{L} = 2$, $P_\mathrm{B} = 0$ dBm, and $R_{\max} = 10$ BPCU.

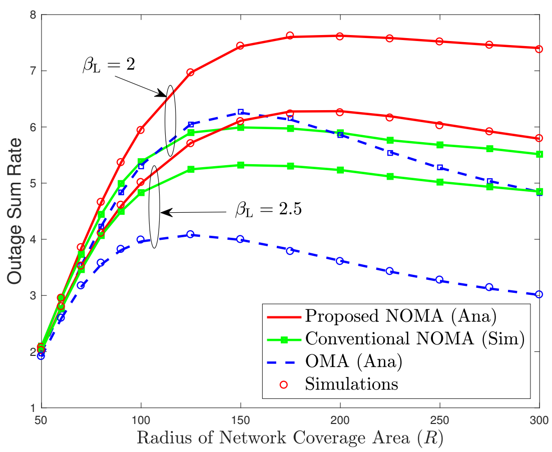


Fig. 6: Outage sum rate versus radius of network coverage area for $\lambda = 0.0001$ nodes/m^2, $\eta = 100$ m, and $P_\mathrm{B} = 0$ dBm.

of NOMA.

In Fig. 5, we show the outage sum rates of the proposed NOMA scheme as a function of the target data rates of both paired NOMA users (i.e., R_S and R_F). As R_S increases, the outage sum rates of all considered schemes improve. Similarly, the outage sum rates of all considered schemes also increase with R_F. However, because the coverage probability decreases with R_S, the rate of improvement of the outage sum rates decreases for large R_S. Moreover, we observe that the outage sum rate gap between NOMA and OMA becomes larger when R_S increases. Hence, the outage sum rate gain of NOMA over OMA is higher when the paired NOMA users have more diverse target data rates. This is because the SINR requirements for OMA to achieve the same spectral efficiency as NOMA increase disproportionally for large R_S.

Fig. 6 shows the impact of the radius of the network coverage area (i.e., R) and the path loss exponent of the LOS links (i.e., β_L) on the outage sum rate. As R increases, the outage sum rate first increases to a peak point as the number of users in the network increases and in turn the probability that two users are served using NOMA increases. However, if R is increased beyond a certain point, the outage sum rate starts to decrease, as both paired NOMA users, especially the typical user, are more likely to suffer from larger path losses. Hence, the size of the network coverage area can be adjusted to maximize spectral efficiency. On the other hand, when β_L is smaller, higher outage sum rates are achieved. This is because signals are less attenuated for smaller path loss exponents, and hence, the network achieves better SINR coverage.

In Fig. 7, we compare the ergodic sum rates of the proposed NOMA scheme as functions of the average LOS range (i.e., η) for different beamsteering error variances (i.e., σ_B^2 and σ_U^2). As η increases, the probabilities that the typical user is LOS and the angle difference $\phi_{0\mathrm{p}}$ is small increase. Hence, the ergodic sum rate of the proposed NOMA scheme increases with η. In addition, the ergodic sum rates of all considered schemes decrease when the beamsteering error variance increases, as the probability that a user is covered by the main lobe of the base station decreases when σ_B^2 and σ_U^2 increase. Moreover, the ergodic sum rate achieved by the proposed NOMA scheme is larger than that achieved by conventional NOMA. This is because, for the proposed user pairing strategy, both paired NOMA users are covered by the main lobe of the base station with a higher probability compared to conventional distance-based user pairing.

Fig. 8 shows the impact of the transmit power (i.e., P_B) and the beamsteering error variances (i.e., σ_B^2 and σ_U^2) on the ergodic sum rate. As P_B increases, the ergodic sum rates of all considered schemes increase, due to a better signal quality. Similar to Fig. 7, the ergodic sum rate of each scheme decreases for higher beamsteering error variances. However, for all considered schemes, the performance difference for different beamsteering error variances decreases with P_B. For the proposed NOMA scheme, there is no performance gap for the two considered values of the beamsteering error variances when $P_\mathrm{B} = 30$ dBm, as in this case, the transmit power is large enough to compensate the negative impact of beam misalignment.

In Fig. 9, we show the ergodic sum rates of the proposed

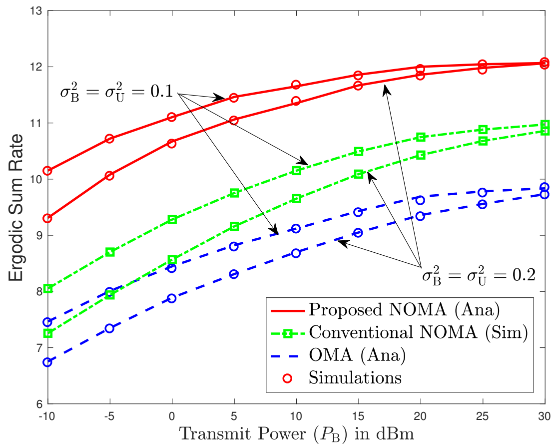


Fig. 8: Ergodic sum rate versus transmit power for different beamsteering error variances, $\lambda = 0.0001$ nodes/m^2, $\beta_L = 2$, $\eta = 100$ m, and $R_{\max} = 10$ BPCU.

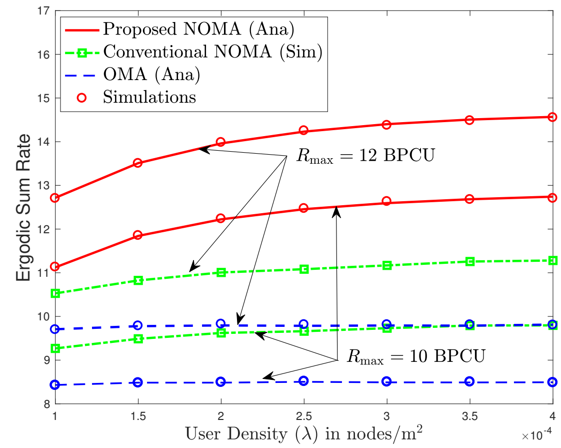


Fig. 10: Ergodic sum rate versus user density of the network for $\beta_L = 2$, $\eta = 100$ m, $P_B = 0$ dBm, and $\sigma_B^2 = \sigma_U^2 = 0.1$.

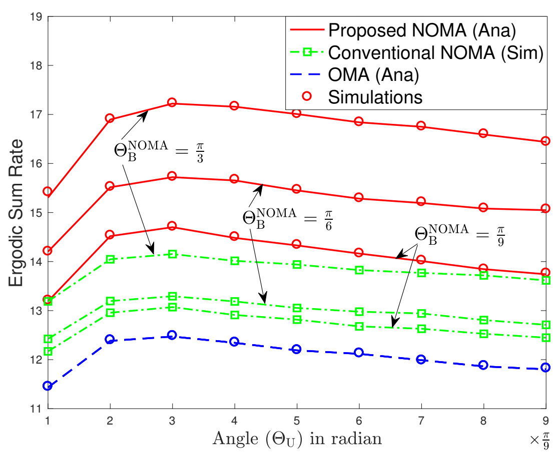


Fig. 9: Ergodic sum rate versus beamwidth of the main lobe when $\lambda = 0.0003$ nodes/m^2, $P_B = 0$ dBm, $\eta = 100$ m, $R_{\max} = 15$ BPCU, and $\sigma_B^2 = \sigma_U^2 = 0.05$.

NOMA scheme as functions of the beamwidth of the main lobe. Considering the impact of Θ_U, there exists a maximum of the ergodic sum rate, due to the tradeoff between the beamwidth of the main lobe and the antenna array gain. Specifically, by increasing Θ_U from $\pi/9$ to $\pi/3$, the ergodic sum rate of the proposed NOMA scheme increases. However, if Θ_U is further increased, the ergodic sum rate of the proposed scheme starts to decrease. This is because, in this case, the benefits introduced by more NOMA transmission opportunities cannot compensate for the reduction in antenna array gain. Furthermore, even for small Θ_B^{NOMA} (e.g., $\Theta_B^{\text{NOMA}} = \pi/9$), the proposed NOMA scheme achieves a much higher ergodic sum rate than OMA. By increasing Θ_B^{NOMA} from $\pi/9$ to $\pi/3$, the ergodic sum rate of the proposed NOMA scheme increases as the probability that NOMA is enabled increases and the NOMA transmission is less vulnerable to beam misalignment.

Fig. 10 illustrates the impact of the user density (i.e., λ) and the maximum achievable rate (i.e., $R_{\max}$) on the ergodic sum rate. As λ increases, the ergodic sum rate of the proposed NOMA scheme increases, as the negative impact of beam misalignment on NOMA transmission decreases. For $\lambda \geq 3.5 \times 10^{-4}$, the ergodic sum rate of the proposed NOMA scheme increases only slowly with λ, saturating to the maximum achievable rate. In contrast, the ergodic sum rate of OMA does not depend on λ. On the other hand, by increasing the maximum possible rate $R_{\max}$ from 10 BPCU to 12 BPCU, the ergodic sum rates of all considered schemes improve, as for larger $R_{\max}$, higher SINRs can be exploited to further enhance the ergodic sum rate. These results show that the maximum achievable rate should be set based on a desired tradeoff between achievable performance and implementation complexity.

VI. Conclusions

In this paper, we developed a general and tractable performance analysis framework for downlink NOMA transmission in mmWave networks with spatially random users taking into account link blockages, directional beamforming, beam misalignment, and user pairing. We proposed an angle-based user pairing strategy, where the LOS user that has the minimum relative angle difference to the typical user is selected. Moreover, we proposed dynamic user ordering among the paired NOMA users to account for the randomness of link blockages and user locations. Tools from stochastic geometry were utilized to derive the coverage probability, outage sum rate, and ergodic sum rate. Simulation results showed that the proposed scheme outperforms conventional NOMA with distance-based user pairing and OMA. Our results revealed that the proposed NOMA scheme reduces the negative impact of beamsteering errors and demonstrated the importance of transmit power allocation and the choice of the beamwidth of the main lobe on the network performance. An interesting topic for future work is the extension of the proposed performance analysis framework to heterogeneous multi-cell networks, where both user association and inter-cell interference need to be taken into account and the impact of the beamwidth on the overall network performance needs to be investigated.

Appendix

A. Proof of Lemma 1

As the spatial locations of the LOS users form an inhomogeneous PPP Φ_{L} with density $\lambda p(r)$, the CDF of the distance between the base station and a randomly selected LOS user is given by

$$F_{\mathrm{L}}(r)=\frac{\int_0^r p(x)x\mathrm{d}x}{\int_0^R p(x)x\mathrm{d}x}\overset{(a)}{=}\frac{1}{\rho}\int_0^r \exp\left(-\frac{x}{\eta}\right)x\mathrm{d}x, \tag{54}$$

where (a) follows by substituting $p(x)=\exp\left(-\frac{x}{\eta}\right)$ and by defining $\rho=\int_0^R \exp(-\frac{x}{\eta})x\mathrm{d}x=\eta^2\left(1-\exp\left(-\frac{R}{\eta}\right)\left(1+\frac{R}{\eta}\right)\right)$. By taking the first derivative of $F_{\mathrm{L}}(r)$ in (54), we obtain the PDF of the distance between the base station and a randomly selected LOS user given in (14). Following similar steps, we obtain the PDF of the distance between the base station and a randomly selected NLOS user, i.e., $f_{\mathrm{N}}(r)$, given in (15).

B. Proof of Lemma 2

For two LOS users U_0 and U_{p}, the distance between the base station and the near LOS user is denoted as $r_{\mathrm{n}}=\min\left(r_0,r_{\mathrm{p}}\right)$. Thereby, the CCDF of distance r_{n} can be expressed as

$$\begin{aligned}\overline{F}_{r_{\mathrm{n}}}(r)&=\mathbb{P}\left(\min\left(r_0,r_{\mathrm{p}}\right)>r\right)\\&\overset{(a)}{=}\mathbb{P}\left(r_0>r\right)\mathbb{P}\left(r_{\mathrm{p}}>r\right)\\&\overset{(b)}{=}\left(1-F_{\mathrm{L}}(r)\right)^2,\end{aligned} \tag{55}$$

where (a) follows from the independence between distances r_0 and r_{p}, and (b) follows by substituting (54) as both users are conditioned to be LOS users and follow the same distance distribution. By taking the first derivative of $1-\overline{F}_{r_{\mathrm{n}}}(r)$, the PDF of distance r_n is given by

$$\begin{aligned}f_{r_{\mathrm{n}}}(r)&=2f_{\mathrm{L}}(r)-2F_{\mathrm{L}}(r)f_{\mathrm{L}}(x)\\&=\frac{2}{\rho}p(r)r-\frac{2}{\rho^2}p(r)r\int_0^r p(x)x\mathrm{d}x.\end{aligned} \tag{56}$$

By substituting $\int_0^r p(x)x\mathrm{d}x=\eta^2\left(1-\exp\left(-\frac{r}{\eta}\right)\left(1+\frac{r}{\eta}\right)\right)$ into (56), we obtain the PDF of distance r_{n} given in (16). On the other hand, the distance between the base station and the far LOS user is denoted as $r_{\mathrm{f}}=\max\left\{r_0,r_{\mathrm{p}}\right\}$. Hence, the CDF of distance r_{f} is given by

$$\begin{aligned}F_{r_{\mathrm{f}}}(r)&=\mathbb{P}\left(\max\left\{r_0,r_{\mathrm{p}}\right\}\le r\right)\\&=\mathbb{P}\left(r_0\le r\right)\mathbb{P}\left(r_{\mathrm{p}}\le r\right)\\&=\left(F_{\mathrm{L}}(r)\right)^2.\end{aligned} \tag{57}$$

By taking the first derivative of $F_{r_{\mathrm{f}}}(r)$, the PDF of distance r_{f} can be derived as

$$f_{r_{\mathrm{f}}}(r)=2F_{\mathrm{L}}(r)f_{\mathrm{L}}(r)=\frac{2}{\rho^2}p(r)r\int_0^r p(x)x\mathrm{d}x. \tag{58}$$

After some algebraic manipulations, we obtain the PDF of distance r_{f} given in (17).

C. Proof of Lemma 3

Although the spatial locations of the LOS users follow an inhomogeneous PPP Φ_{L}, the spatial distribution of the LOS users is isotropic. As angle ϕ_i is uniformly distributed within $[0,2\pi]$ and is independent of angle $\phi_j, j\neq i$, the CDF of $\phi_{0i}=|\phi_0-\phi_i|$ can be expressed as $F_{\phi_{0i}}(\phi)=\frac{4\pi\phi-\phi^2}{4\pi^2}, \phi\in[0,2\pi]$. By taking the first derivative of $F_{\phi_{0i}}(\phi)$, the PDF of ϕ_{0i} is given by $f_{\phi_{0i}}(\phi)=\frac{\mathrm{d}}{\mathrm{d}\phi}F_{\phi_{0i}}(\phi)=\frac{2\pi-\phi}{2\pi^2}, \phi\in[0,2\pi]$. When there are k_{L} LOS users in the network coverage area and user U_0 is LOS, the CDF of the minimum relative angle difference $\phi_{0\mathrm{p}}$ can be expressed as

$$F_{\phi_{0\mathrm{p}}}(\phi)=1-\left(\frac{2\pi-\phi}{2\pi}\right)^{2k_{\mathrm{L}}-2}, \tag{59}$$

which follows from the user pairing strategy in (3) and by applying order statistics [41]. When there are k_{L} LOS users in the network coverage area but user U_0 is NLOS, by following similar steps, we obtain the CDF of the minimum relative angle difference $\phi_{0\mathrm{p}}$ as

$$F_{\phi_{0\mathrm{p}}}(\phi)=1-\left(\frac{2\pi-\phi}{2\pi}\right)^{2k_{\mathrm{L}}}. \tag{60}$$

By taking the first derivative of $F_{\phi_{0\mathrm{p}}}(\phi)$ given in (59) and (60), we obtain the PDF of angle difference $\phi_{0\mathrm{p}}$ in (18).

D. Proof of Lemma 4

When the angle difference $\phi_{0\mathrm{p}}$ is given, typical user U_0 is covered by the main lobe of the base station if $|\frac{1}{2}\phi_{0\mathrm{p}}+\Delta_{\mathrm{B}}|\le\frac{1}{2}\Theta_{\mathrm{B}}^{\mathrm{NOMA}}$, which can equivalently be expressed as $\Delta_{\mathrm{B}}\in\left[-\frac{1}{2}\left(\Theta_{\mathrm{B}}^{\mathrm{NOMA}}+\phi_{0\mathrm{p}}\right),\frac{1}{2}\left(\Theta_{\mathrm{B}}^{\mathrm{NOMA}}-\phi_{0\mathrm{p}}\right)\right]$. As beamsteering error Δ_{B} is assumed to follow a Gaussian distribution, the probabilities that typical user U_0 is covered by the main lobe and the side lobe of the base station can be expressed as $g(\phi_{0\mathrm{p}})=F_{\Delta_{\mathrm{B}}}\left(\frac{\Theta_{\mathrm{B}}^{\mathrm{NOMA}}-\phi_{0\mathrm{p}}}{2}\right)-F_{\Delta_{\mathrm{B}}}\left(-\frac{\Theta_{\mathrm{B}}^{\mathrm{NOMA}}+\phi_{0\mathrm{p}}}{2}\right)$ and $1-g(\phi_{0\mathrm{p}})$, respectively. Hence, the PMF of the directivity gain from the base station to the typical user can be expressed as $\mathbb{P}\left(G_{\mathrm{B}}\left(\Theta_{\mathrm{B}}^{\mathrm{NOMA}},\frac{1}{2}\phi_{0\mathrm{p}}+\Delta_{\mathrm{B}}\right)=G_{\mathrm{B}}^{\mathrm{M}}\left(\Theta_{\mathrm{B}}^{\mathrm{NOMA}}\right)\right)=g(\phi_{0\mathrm{p}})$ and $\mathbb{P}\left(G_{\mathrm{B}}\left(\Theta_{\mathrm{B}}^{\mathrm{NOMA}},\frac{1}{2}\phi_{0\mathrm{p}}+\Delta_{\mathrm{B}}\right)=G_{\mathrm{B}}^{\mathrm{S}}\left(\Theta_{\mathrm{B}}^{\mathrm{NOMA}}\right)\right)=1-g(\phi_{0\mathrm{p}})$. The base station is covered by the main lobe of typical user U_0 if beamsteering error $\Delta_{\mathrm{U}_0}\in\left[-\frac{1}{2}\Theta_{\mathrm{U}},\frac{1}{2}\Theta_{\mathrm{U}}\right]$. Hence, the probability that the base station is covered by the main lobe and the side lobe of the typical user can be expressed as $y=F_{\Delta_{\mathrm{U}}}\left(\frac{1}{2}\Theta_{\mathrm{U}}\right)-F_{\Delta_{\mathrm{U}}}\left(-\frac{1}{2}\Theta_{\mathrm{U}}\right)$ and $1-y$, respectively. Thereby, the PMF of the directivity gain from the typical user to the base station can be expressed as $\mathbb{P}\left(G_{\mathrm{U}}\left(\Theta_{\mathrm{U}},\Delta_{\mathrm{U}_0}\right)=G_{\mathrm{U}}^{\mathrm{M}}\left(\Theta_{\mathrm{U}}\right)\right)=y$ and $\mathbb{P}\left(G_{\mathrm{U}}\left(\Theta_{\mathrm{U}},\Delta_{\mathrm{U}_0}\right)=G_{\mathrm{U}}^{\mathrm{S}}\left(\Theta_{\mathrm{U}}\right)\right)=1-y$.

As $\mathcal{D}_0^{\mathrm{NOMA}}=G_{\mathrm{B}}\left(\Theta_{\mathrm{B}}^{\mathrm{NOMA}},\frac{1}{2}\phi_{0\mathrm{p}}+\Delta_{\mathrm{B}}\right)G_{\mathrm{U}}\left(\Theta_{\mathrm{U}},\Delta_{\mathrm{U}_0}\right)$, the PMF of the total directivity gain between the base station and the typical user when NOMA is enabled is given by

$$\begin{aligned}\mathbb{P}\left(\mathcal{D}_0^{\mathrm{NOMA}}=\vartheta\right)=&\sum_{x\in\Lambda}\mathbb{P}\left(G_{\mathrm{B}}\left(\Theta_{\mathrm{B}}^{\mathrm{NOMA}},\frac{1}{2}\phi_{0\mathrm{p}}+\Delta_{\mathrm{B}}\right)=x\right)\\&\times\mathbb{P}\left(G_{\mathrm{U}}\left(\Theta_{\mathrm{U}},\Delta_{\mathrm{U}_0}\right)=\frac{\vartheta}{x}\right),\end{aligned} \tag{61}$$

where $\Lambda=\left\{G_{\mathrm{B}}^{\mathrm{M}}\left(\Theta_{\mathrm{B}}^{\mathrm{NOMA}}\right),G_{\mathrm{B}}^{\mathrm{S}}\left(\Theta_{\mathrm{B}}^{\mathrm{NOMA}}\right)\right\}$.

By substituting the PMFs of directivity gains $G_{\mathrm{B}}\left(\Theta_{\mathrm{B}}^{\mathrm{NOMA}}, \frac{1}{2}\phi_{0\mathrm{p}}+\Delta_{\mathrm{B}}\right)$ and $G_{\mathrm{U}}\left(\Theta_{\mathrm{U}}, \Delta_{\mathrm{U}_0}\right)$ into (61), we obtain the PMF of the total directivity gain $\mathcal{D}_0^{\mathrm{NOMA}}$ given in (19). Following similar steps, we obtain the PMF of the total directivity gain $\mathcal{D}_{\mathrm{p}}^{\mathrm{NOMA}}$ given in (20).

E. Proof of Lemma 5

When OMA is enabled, typical user U_0 is covered by the main lobe of the base station if $\Delta_{\mathrm{B}} \in [-\frac{1}{2}\Theta_{\mathrm{B}}^{\mathrm{OMA}}, \frac{1}{2}\Theta_{\mathrm{B}}^{\mathrm{OMA}}]$. Hence, the probabilities that the typical user is covered by the main lobe and the side lobe of the base station are given by $w = F_{\Delta_{\mathrm{B}}}\left(\frac{1}{2}\Theta_{\mathrm{B}}^{\mathrm{OMA}}\right) - F_{\Delta_{\mathrm{B}}}\left(-\frac{1}{2}\Theta_{\mathrm{B}}^{\mathrm{OMA}}\right)$ and $1-w$, respectively. The PMF of the directivity antenna gain from the base station to the typical user when OMA is enabled can be expressed as $\mathbb{P}\left(G_{\mathrm{B}}\left(\Theta_{\mathrm{B}}^{\mathrm{OMA}}, \Delta_{\mathrm{B}}\right) = G_{\mathrm{B}}^{\mathrm{M}}(\Theta_{\mathrm{B}}^{\mathrm{OMA}})\right) = w$ and $\mathbb{P}\left(G_{\mathrm{B}}\left(\Theta_{\mathrm{B}}^{\mathrm{OMA}}, \Delta_{\mathrm{B}}\right) = G_{\mathrm{B}}^{\mathrm{S}}(\Theta_{\mathrm{B}}^{\mathrm{OMA}})\right) = 1-w$. Following similar steps as in the proof of Lemma 4, the PMF of the total directivity gain between the base station and the typical user for OMA transmission given in (24) is obtained.

F. Proof of Proposition 1

Given that there are $k_{\mathrm{L}} \geq 2$ LOS users, the probability that user U_0 can successfully perform SIC and decode its own signal is given by

$$
\begin{aligned}
Q_{\mathrm{I},0}(\xi_1) &= \mathbb{P}\left(\gamma_{\mathrm{p}\to 0} \geq T_{\mathrm{F}}, \gamma_0 \geq T_{\mathrm{S}}, \phi_{0\mathrm{p}} \leq \Theta_{\mathrm{B}}^{\mathrm{NOMA}}, r_0 \leq r_{\mathrm{p}}\right) \\
&= \int_0^{\Theta_{\mathrm{B}}^{\mathrm{NOMA}}} \mathbb{P}\left(\gamma_{\mathrm{p}\to 0} \geq T_{\mathrm{F}}, \gamma_0 \geq T_{\mathrm{S}}, r_0 \leq r_{\mathrm{p}}\right) f_{\phi_{0\mathrm{p}}}(\phi_{0\mathrm{p}})\mathrm{d}\phi_{0\mathrm{p}}, \quad (62)
\end{aligned}
$$

where $f_{\phi_{0\mathrm{p}}}(\phi_{0\mathrm{p}})$ is given in (18). Given the angle difference $\phi_{0\mathrm{p}}$, we have

$$
\begin{aligned}
&\mathbb{P}\left(\gamma_{\mathrm{p}\to 0} \geq T_{\mathrm{F}}, \gamma_0 \geq T_{\mathrm{S}}, r_0 \leq r_{\mathrm{p}}\right) \\
&\overset{(a)}{=} \mathbb{P}\left(|h_0|^2 \geq \frac{\xi_1}{\mathcal{D}_0^{\mathrm{NOMA}} \ell_{\mathrm{L}}(r_{\mathrm{n}})}\right) \\
&\overset{(b)}{=} \sum_{n=0}^{N_{\mathrm{L}}-1} \frac{(N_{\mathrm{L}}\xi_1)^n}{n!} \mathbb{E}_{\{\mathcal{D}_0^{\mathrm{NOMA}}, r_{\mathrm{n}}\}}\left[\frac{\exp\left(-\frac{N_{\mathrm{L}}\xi_1}{\mathcal{D}_0^{\mathrm{NOMA}}\ell_{\mathrm{L}}(r_{\mathrm{n}})}\right)}{\left(\mathcal{D}_0^{\mathrm{NOMA}}\ell_{\mathrm{L}}\left(r_{\mathrm{n}}\right)\right)^n}\right], \quad (63)
\end{aligned}
$$

where (a) follows by substituting (4) and (6) and by defining $\xi_1 = \max\left\{\frac{T_{\mathrm{F}}\sigma^2}{(\alpha_{\mathrm{F}}^2 - T_{\mathrm{F}}\alpha_{\mathrm{S}}^2)P_{\mathrm{B}}}, \frac{T_{\mathrm{S}}\sigma^2}{\alpha_{\mathrm{S}}^2 P_{\mathrm{B}}}\right\}$, and (b) follows from the normalized Gamma distribution of $|h_0|^2$. Note that $\mathbb{E}_{\{\mathcal{D}_0^{\mathrm{NOMA}}, r_{\mathrm{n}}\}}[\cdot]$ refers to the expectation over the total directivity gain $\mathcal{D}_0^{\mathrm{NOMA}}$ and distance r_{n}, which takes into account that typical user U_0 is the near LOS user, i.e., $r_0 \leq r_{\mathrm{p}}$.

By utilizing the PMF of the total directivity gain $\mathcal{D}_0^{\mathrm{NOMA}}$ given in Lemma 4, we have

$$
\begin{aligned}
&\mathbb{E}_{\{\mathcal{D}_0^{\mathrm{NOMA}}, r_{\mathrm{n}}\}}\left[\frac{\exp\left(-\frac{N_{\mathrm{L}}\xi_1}{\mathcal{D}_0^{\mathrm{NOMA}}\ell_{\mathrm{L}}(r_{\mathrm{n}})}\right)}{\left(\mathcal{D}_0^{\mathrm{NOMA}}\ell_{\mathrm{L}}\left(r_{\mathrm{n}}\right)\right)^n}\right] \\
&\overset{(a)}{=} \mathbb{E}_{\{r_{\mathrm{n}}\}}\left[\sum_{i=1}^{4} \frac{d_i(\phi_{0\mathrm{p}})}{(c_i\ell_{\mathrm{L}}(r_{\mathrm{n}}))^n} \exp\left(-\frac{N_{\mathrm{L}}\xi_1}{c_i\ell_{\mathrm{L}}(r_{\mathrm{n}})}\right)\right] \\
&\overset{(b)}{=} \sum_{i=1}^{4} d_i(\phi_{0\mathrm{p}}) \int_0^R \left(\frac{r_{\mathrm{n}}^{\beta_{\mathrm{L}}}}{c_i C_{\mathrm{L}}}\right)^n \exp\left(-\frac{N_{\mathrm{L}}\xi_1 r_{\mathrm{n}}^{\beta_{\mathrm{L}}}}{c_i C_{\mathrm{L}}}\right) f_{r_{\mathrm{n}}}(r_{\mathrm{n}})\mathrm{d}r_{\mathrm{n}} \\
&\overset{(c)}{=} \sum_{i=1}^{4} d_i(\phi_{0\mathrm{p}}) \frac{R\pi}{2M_2} \sum_{m_2=1}^{M_2} \sqrt{1-\varpi_{m_2}^2} \left(\frac{r_{\mathrm{n},m_2}^{\beta_{\mathrm{L}}}}{c_i C_{\mathrm{L}}}\right)^n \\
&\quad \times \exp\left(-\frac{N_{\mathrm{L}}\xi_1 r_{\mathrm{n},m_2}^{\beta_{\mathrm{L}}}}{c_i C_{\mathrm{L}}}\right) f_{r_{\mathrm{n}}}(r_{\mathrm{n},m_2}), \quad (64)
\end{aligned}
$$

where $r_{\mathrm{n},m_2} = \frac{R}{2}(\varpi_{m_2}+1)$, $\varpi_{m_2} = \cos\left(\frac{2m_2-1}{2M_2}\pi\right)$, (a) and (b) follow from the expectations over $\mathcal{D}_0^{\mathrm{NOMA}}$ and r_{n}, respectively, and (c) follows by applying Gauss-Chebyshev quadrature [38]. By substituting (63) and (64) into (62) and by applying Gauss-Chebyshev quadrature with respect to the integral over $\phi_{0\mathrm{p}}$, we obtain $Q_{\mathrm{I},0}(\xi_1)$ in (27). Considering the randomness of the number of the LOS and NLOS users, we obtain (26) as there are at least two LOS users in this case.

G. Proof of Corollary 1

Without beamsteering errors, the base station and user U_0 are always covered by each other's main lobe when $\phi_{0\mathrm{p}} \leq \Theta_{\mathrm{B}}^{\mathrm{NOMA}}$. As a result, the total directivity gain is given by $\mathcal{D}_0^{\mathrm{NOMA}} = G_{\mathrm{B}}^{\mathrm{M}}(\Theta_{\mathrm{B}}^{\mathrm{NOMA}})G_{\mathrm{U}}^{\mathrm{M}}(\Theta_{\mathrm{U}}) = c_1$ and the probability of $\phi_{0\mathrm{p}} \leq \Theta_{\mathrm{B}}^{\mathrm{NOMA}}$ is $1-\left(1-\frac{\Theta_{\mathrm{B}}^{\mathrm{NOMA}}}{2\pi}\right)^{2k_{\mathrm{L}}-2}$. For the LOS ball model, the CDF of the distance between the base station and a randomly selected LOS user is given by $F_{\mathrm{L}}(r) = \frac{r^2}{R_{\mathrm{LOS}}^2}\mathbb{1}(r < R_{\mathrm{LOS}})$. By taking the first derivative of $F_{\mathrm{L}}(r)$, we obtain the PDF as $f_{\mathrm{L}}(r) = \frac{2r}{R_{\mathrm{LOS}}^2}\mathbb{1}(r < R_{\mathrm{LOS}})$. By following similar steps as in the proof of Lemma 2, the PDF of distance r_n is obtained as $f_{r_{\mathrm{n}}}(r) = 2f_{\mathrm{L}}(r) - 2F_{\mathrm{L}}(r)f_{\mathrm{L}}(r) = \frac{4r}{R_{\mathrm{LOS}}^2} - \frac{4r^3}{R_{\mathrm{LOS}}^4}$, for $r < R_{\mathrm{LOS}}$. By substituting $f_{r_{\mathrm{n}}}(r)$ into the following integral, we have

$$
\begin{aligned}
&\int_0^R r^{n\beta_{\mathrm{L}}} \exp\left(-\frac{N_{\mathrm{L}}\xi_1}{c_1 C_{\mathrm{L}}} r^{\beta_{\mathrm{L}}}\right) f_{r_{\mathrm{n}}}(r)\mathrm{d}r \\
&= \frac{4}{\beta_{\mathrm{L}} R_{\mathrm{LOS}}^2}\left(\frac{c_1 C_{\mathrm{L}}}{N_{\mathrm{L}}\xi_1}\right)^{n+\frac{1}{\beta_{\mathrm{L}}}} \gamma\left(n+\frac{1}{\beta_{\mathrm{L}}}, \frac{N_{\mathrm{L}}\xi_1}{c_1 C_{\mathrm{L}}} R_{\mathrm{LOS}}^{\beta_{\mathrm{L}}}\right) \\
&\quad - \frac{4}{\beta_{\mathrm{L}} R_{\mathrm{LOS}}^4}\left(\frac{c_1 C_{\mathrm{L}}}{N_{\mathrm{L}}\xi_1}\right)^{n+\frac{4}{\beta_{\mathrm{L}}}} \gamma\left(n+\frac{4}{\beta_{\mathrm{L}}}, \frac{N_{\mathrm{L}}\xi_1}{c_1 C_{\mathrm{L}}} R_{\mathrm{LOS}}^{\beta_{\mathrm{L}}}\right).
\end{aligned}
$$

After combining the aforementioned results, we obtain $Q_{\mathrm{I},0}(\xi_1)$ given in (29).

H. Proof of Proposition 2

Given that there are $k_{\mathrm{L}} \geq 2$ LOS users, the probability that user U_{p} can successfully decode its own signal is given by

$$
\begin{aligned}
Q_{\mathrm{I,p}}(\xi_2) &= \mathbb{P}\left(\gamma_{\mathrm{p}|0} \geq T_{\mathrm{F}}, \phi_{0\mathrm{p}} \leq \Theta_{\mathrm{B}}^{\mathrm{NOMA}}, r_{\mathrm{p}} \geq r_0\right) \\
&= \int_0^{\Theta_{\mathrm{B}}^{\mathrm{NOMA}}} \mathbb{P}\left(\gamma_{\mathrm{p}|0} \geq T_{\mathrm{F}}, r_{\mathrm{p}} \geq r_0\right) f_{\phi_{0\mathrm{p}}}(\phi_{0\mathrm{p}})\mathrm{d}\phi_{0\mathrm{p}}. \quad (65)
\end{aligned}
$$

Following similar steps as in the proof of Proposition 1, we obtain $Q_{\mathrm{I,p}}(\xi_2)$ given in (31) and $P_{\mathrm{I,p}}^{\mathrm{cov}}(\xi_2)$ given in (30), by taking into account that paired NOMA user U_{p} is the far LOS user (i.e., $r_{\mathrm{p}} \geq r_0$) and the PMF of the total directivity gain $\mathcal{D}_{\mathrm{p}}^{\mathrm{NOMA}}$ is given in (20).

I. Proof of Proposition 3

In NOMA-I, the CCDF of γ_0 defined in (6) can be expressed as

$$\overline{F}_{\gamma_0}(x) = \sum_{k_{\mathrm{L}}=2}^{\infty}\sum_{k_{\mathrm{N}}=0}^{\infty} \frac{k_{\mathrm{L}}}{k_{\mathrm{L}}+k_{\mathrm{N}}}\Psi_{\mathrm{L}}(k_{\mathrm{L}})\Psi_{\mathrm{N}}(k_{\mathrm{N}}) \times \int_0^{\Theta_{\mathrm{B}}^{\mathrm{NOMA}}}\int_0^{R} \mathbb{P}(\gamma_0 > x) f_{r_{\mathrm{n}}}(r_{\mathrm{n}})\mathrm{d}r_{\mathrm{n}} f_{\phi_{0\mathrm{p}}}(\phi_{0\mathrm{p}})\mathrm{d}\phi_{0\mathrm{p}}, \quad (66)$$

where $\mathbb{P}\left(\gamma_0 > x\right) = \mathbb{P}\left(|h_0|^2 > \frac{\xi_3(x)}{\mathcal{D}_0^{\mathrm{NOMA}}\ell_{\mathrm{L}}(r_{\mathrm{n}})}\right)$. By following the same steps as in the proof of Proposition 1, we obtain $\overline{F}_{\gamma_0}(x) = P_{\mathrm{I},0}(\xi_3(x))$, where $\xi_3(x) = \frac{\sigma^2 x}{\alpha_{\mathrm{S}}^2 P_{\mathrm{B}}}$. Similarly, in NOMA-I, the CCDF of $\gamma_{\mathrm{p}|0}$ given in (7) can be computed by $\overline{F}_{\gamma_{\mathrm{p}|0}}(x) = P_{\mathrm{I,p}}(\xi_4(x))$, where $\xi_4(x) = \frac{\sigma^2 x}{(\alpha_{\mathrm{F}}^2 - x\alpha_{\mathrm{S}}^2)P_{\mathrm{B}}}$. Note that when $x > \frac{\alpha_{\mathrm{F}}^2}{\alpha_{\mathrm{S}}^2}$, we have $\overline{F}_{\gamma_{\mathrm{p}|0}}(x) = 0$. Thus, the ergodic sum rate of users U_0 and U_{p} is given by

$$\mathcal{R}_{\mathrm{I}}^{\mathrm{NOMA}} = \frac{1}{\ln 2}\int_0^{T_{\max}} \frac{P_{\mathrm{I},0}(\xi_3(x))}{1+x}\mathrm{d}x + \frac{1}{\ln 2}\int_0^{\min\left\{T_{\max}, \frac{\alpha_{\mathrm{F}}^2}{\alpha_{\mathrm{S}}^2}\right\}} \frac{P_{\mathrm{I,p}}(\xi_4(x))}{1+x}\mathrm{d}x. \quad (67)$$

By applying Gauss-Chebyshev quadrature, we obtain $\mathcal{R}_{\mathrm{I}}^{\mathrm{NOMA}}$ given in (49).

References

[1] Y. Zhou, V. W.S. Wong, and R. Schober, "Performance analysis of millimeter wave NOMA networks with beam misalignment," in *Proc. of IEEE Int'l Conf. on Commun. (ICC)*, Kansas, MO, May 2018.

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

[3] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, and H. V. Poor, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 185–191, Feb. 2017.

[4] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *Proc. of IEEE PIMRC*, London, UK, Sept. 2013.

[5] Z. Ding, P. Fan, and V. Poor, "Impact of user pairing on 5G non-orthogonal multiple access downlink transmissions," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6010–6023, Sept. 2016.

[6] J. Choi, "On the power allocation for MIMO-NOMA systems with layered transmissions," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3226–3237, May 2016.

[7] 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.

[8] Y. Zhou and V. W.S. Wong, "Stable throughput region of downlink NOMA transmissions with limited CSI," in *Proc. of IEEE Int'l Conf. on Commun. (ICC)*, Paris, France, May 2017.

[9] Y. Zhou, V. W.S. Wong, and R. Schober, "Stable throughput regions of opportunistic NOMA and cooperative NOMA with full-duplex relaying," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 5059–5075, Aug. 2018.

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

[11] Y. Zhou, V. W.S. Wong, and R. Schober, "Performance analysis of cooperative NOMA with dynamic decode-and-forward relaying," in *Proc. of IEEE Globecom*, Singapore, Dec. 2017.

[12] ——, "Dynamic decode-and-forward based cooperative NOMA with spatially random users," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3340–3356, May 2018.

[13] W. Roh, J.-Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 106–113, Feb. 2014.

[14] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. of the IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.

[15] S. Mumtaz, J. Rodriguez, and L. Dai, *mmWave Massive MIMO: A Paradigm for 5G*. Academic Press, 2016.

[16] 3GPP TR 36.873, "Technical specification group radio access network: Study on 3D channel model for LTE," Sept. 2017.

[17] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.

[18] T. Bai and R. W. Heath, "Coverage and rate analysis for millimeter-wave cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 1100–1114, Feb. 2015.

[19] M. Di Renzo, "Stochastic geometry modeling and analysis of multi-tier millimeter wave cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 9, pp. 5038–5057, Sep. 2015.

[20] X. Yu, J. Zhang, M. Haenggi, and K. B. Letaief, "Coverage analysis for millimeter wave networks: The impact of directional antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1498–1512, Jul. 2017.

[21] A. Tassi, M. Egan, R. J. Piechocki, and A. Nix, "Modeling and design of millimeter-wave networks for highway vehicular communication," *IEEE Trans. Veh. Technol.*, vol. 66, no. 12, pp. 10 676–10 691, Dec. 2017.

[22] M. Rebato, J. Park, P. Popovski, E. De Carvalho, and M. Zorzi, "Stochastic geometric coverage analysis in mmWave cellular networks with a realistic channel model," in *Proc. of IEEE Globecom*, Singapore, Dec. 2017.

[23] A. Alkhateeb, Y.-H. Nam, M. S. Rahman, J. Zhang, and R. W. Heath, "Initial beam association in millimeter wave cellular systems: Analysis and design insights," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2807–2821, May 2017.

[24] 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.

[25] Z. Ding, P. Fan, and H. Poor, "Random beamforming in millimeter-wave NOMA networks," *IEEE Access*, vol. 5, pp. 7667–7681, Feb. 2017.

[26] Z. Ding, L. Dai, R. Schober, and H. V. Poor, "NOMA meets finite resolution analog beamforming in massive MIMO and millimeter-wave networks," *IEEE Commun. Lett.*, no. 8, pp. 1879–1882, Aug. 2017.

[27] D. Zhang, Z. Zhou, C. Xu, Y. Zhang, J. Rodriguez, and T. Sato, "Capacity analysis of NOMA with mmWave massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1606–1618, Jul. 2017.

[28] B. Wang, L. Dai, Z. Wang, N. Ge, and S. Zhou, "Spectrum and energy-efficient beamspace MIMO-NOMA for millimeter-wave communications using lens antenna array," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2370–2382, Oct. 2017.

[29] J. Wildman, P. H. J. Nardelli, M. Latva-aho, and S. Weber, "On the joint impact of beamwidth and orientation error on throughput in directional wireless Poisson networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 12, pp. 7072–7085, Dec. 2014.

[30] T. Rappaport, R. Heath, R. Daniels, and J. Murdock, *Millimeter Wave Wireless Communications*. Pearson PLC, 2014.

[31] T. Bai, R. Vaze, and R. W. Heath, "Analysis of blockage effects on urban cellular networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 5070–5083, Sep. 2014.

[32] J. G. Andrews, T. Bai, M. N. Kulkarni, A. Alkhateeb, A. K. Gupta, and R. W. Heath, "Modeling and analyzing millimeter wave cellular systems," *IEEE Trans. Commun.*, vol. 65, no. 1, pp. 403–430, Jan. 2017.

[33] F. Baccelli and B. Błaszczyszyn, *Stochastic Geometry and Wireless Networks - Part I: Theory*. Foundations and Trends in Networking, Now Publishers, Nov. 2010.

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

[35] S. Singh, M. N. Kulkarni, A. Ghosh, and J. G. Andrews, "Tractable model for rate in self-backhauled millimeter wave cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 10, pp. 2196–2211, Oct. 2015.

[36] 3GPP TR 36.859, "Study on downlink multiuser superposition transmission (MUST) for LTE," Jan. 2016.

[37] 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.

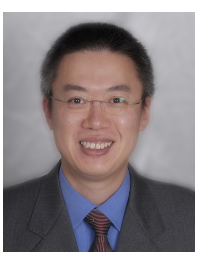
[38] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover Publications, 1972.

[39] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3122–3134, Nov. 2011.

[40] 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.

[41] H. A. David and H. N. Nagaraja, *Order Statistics*. John Wiley, New York, 2003.

Yong Zhou (S'13, M'16) received the B.Sc. and M.Eng. degrees from Shandong University, Jinan, China, in 2008 and 2011, respectively, and the Ph.D. degree from the University of Waterloo, Waterloo, ON, Canada, in 2015. From 2015 to 2017, he worked as a postdoctoral research fellow in the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, Canada. He is currently an Assistant Professor in the School of Information Science and Technology, ShanghaiTech University, China. His research interests include performance analysis and resource allocation of 5G and Internet of Things (IoT) networks. He has served as a technical program committee (TPC) member for several conferences.



Vincent W.S. Wong (S'94, M'00, SM'07, F'16) received the B.Sc. degree from the University of Manitoba, Winnipeg, MB, Canada, in 1994, the M.A.Sc. degree from the University of Waterloo, Waterloo, ON, Canada, in 1996, and the Ph.D. degree from the University of British Columbia (UBC), Vancouver, BC, Canada, in 2000. From 2000 to 2001, he worked as a systems engineer at PMC-Sierra Inc. (now Microchip Technology Inc.). He joined the Department of Electrical and Computer Engineering at UBC in 2002 and is currently a Professor. His research areas include protocol design, optimization, and resource management of communication networks, with applications to wireless networks, smart grid, mobile edge computing, and Internet of Things. Currently, Dr. Wong is an Executive Editorial Committee Member of the *IEEE Transactions on Wireless Communications*, an Editor-at-Large of the *IEEE Transactions on Communications*, and an Associate Editor of the *IEEE Transactions on Mobile Computing*. He has served as a Guest Editor of *IEEE Journal on Selected Areas in Communications* and *IEEE Wireless Communications*. He has also served on the editorial boards of *IEEE Transactions on Vehicular Technology* and *Journal of Communications and Networks*. He is a Tutorial Co-Chair of *IEEE Globecom*'18. He was a Technical Program Co-chair of *IEEE SmartGridComm*'14, as well as a Symposium Co-chair of *IEEE ICC*'18, *IEEE SmartGridComm* ('13, '17) and *IEEE Globecom*'13. He is the Chair of the IEEE Vancouver Joint Communications Chapter and has served as the Chair of the IEEE Communications Society Emerging Technical Sub-Committee on Smart Grid Communications. He received the 2014 UBC Killam Faculty Research Fellowship.

Robert Schober (S'98, M'01, SM'08, F'10) received the Diplom (Univ.) and the Ph.D. degrees in electrical engineering from Friedrich-Alexander University of Erlangen-Nuremberg (FAU), Germany, in 1997 and 2000, respectively. From 2002 to 2011, he was a Professor and Canada Research Chair at the University of British Columbia (UBC), Vancouver, Canada. Since January 2012 he is an Alexander von Humboldt Professor and the Chair for Digital Communication at FAU. His research interests fall into the broad areas of Communication Theory, Wireless Communications, and Statistical Signal Processing.

Robert received several awards for his work including the 2002 Heinz MaierLeibnitz Award of the German Science Foundation (DFG), the 2004 Innovations Award of the Vodafone Foundation for Research in Mobile Communications, a 2006 UBC Killam Research Prize, a 2007 Wilhelm Friedrich Bessel Research Award of the Alexander von Humboldt Foundation, the 2008 Charles McDowell Award for Excellence in Research from UBC, a 2011 Alexander von Humboldt Professorship, a 2012 NSERC E.W.R. Stacie Fellowship, and a 2017 Wireless Communications Recognition Award by the IEEE Wireless Communications Technical Committee. He is listed as a 2017 Highly Cited Researcher by the Web of Science and a Distinguished Lecturer of the IEEE Communications Society (ComSoc). Robert is a Fellow of the Canadian Academy of Engineering and a Fellow of the Engineering Institute of Canada. From 2012 to 2015, he served as Editor-in-Chief of the IEEE Transactions on Communications. Currently, he is the Chair of the Steering Committee of the IEEE Transactions on Molecular, Biological and Multiscale Communication, a Member of the Editorial Board of the Proceedings of the IEEE, a Member at Large of the Board of Governors of ComSoc, and the ComSoc Director of Journals.