The Effectiveness of Hybrid Beamforming in Enhancing the Performance of NOMA-mmWave and Massive MIMO Systems

Mourtada OUBASSGHIR[1,*] and *Mohamed* BOULOUIRD[2]

[1]SSA Group, ENSA-M, Cadi Ayyad University Marrakesh, Morocco
[2]SSA Group, ENSA-M, Cadi Ayyad University Marrakesh, Morocco

Abstract. As wireless networks evolve, the integration of Non-Orthogonal Multiple Access (NOMA), Massive Multiple Input Multiple Output (mMIMO), and millimeter wave (mmWave) technologies emerges as an effective approach to drive this advancement. This integration boosts both spectral efficiency (SE) and energy efficiency (EE), ensuring more reliable, faster, and energy-efficient communication systems for wireless communications, including 5G and beyond. This study aims to examine the performance of this integration in terms of SE and EE by employing a system model featuring a single base station equipped with multiple antennas that generate distinct beams serving multiple users using max-min fairness power allocation. We compare the achieved SE and EE with the Orthogonal Multiple Access (OMA) technologies used in previous generations. Furthermore, we explore the effect of using hybrid and full digital beamforming architectures on energy consumption. Through this analysis, we aim to demonstrate how employing NOMA-mMIMO and mmWave integration with a hybrid structure can achieve high SE and low energy consumption.

1 Introduction

Wireless networks have continuously evolved, with each decade bringing more advanced generations [1]. The core goals of these developments have been to increase data rates and reduce latency [2, 3]. A pivotal element in this evolution is the progression of multiple access technologies. This began with Frequency Division Multiple Access (FDMA) in the first generation (1G), which allocated different frequency bands to different users. The second generation (2G) followed with Time Division Multiple Access (TDMA), dividing each frequency band into time slots for multiple users. The third generation (3G) introduced Code Division Multiple Access (CDMA), allowing multiple users to share the same frequency band by assigning unique codes to each signal. The fourth generation (4G) brought Orthogonal Frequency Division Multiple Access (OFDMA), enhancing efficiency by dividing the spectrum into small frequency bands, known as subcarriers, to minimize interference [1].

Despite these remarkable advancements, the increasing demand for mobile services has underscored the limitations of existing technologies, notably their inefficient spectrum usage. The introduction of NOMA marked a significant paradigm shift. Unlike traditional methods that separate users by time, frequency, or code [4, 5], NOMA allows multiple users to share the same frequencies, distinguishing them through power levels and employing sophisticated signal processing techniques. This innovation promises to dramatically optimize spectrum usage, accommodate more users, and facilitate higher data rates [6]. On the transmission side, NOMA employs the superposition coding technique, wherein the base station (BS) overlays user signals at varying power levels within the same frequency band. This power allocation is tailored to users' channel conditions, assigning more power to users with weaker channel conditions and less to those with stronger ones. On the receiving side, users utilize successive interference cancellation (SIC) to eliminate interference, effectively decoding their own signal from the combined ones [7].

Furthermore, mmWave refers to electromagnetic waves with frequencies between 30 to 300 GHz, matching to wavelengths of 1-10 mm, however, the 24. 25-29.5 GHz band is widely used for 5G New Radio [8]. This offers a broad bandwidth capable of delivering data at speeds of several gigabits per second. The higher frequencies lead to shorter wavelengths, which allows for the use of smaller antennas. Consequently, this increases the possibility of using a BS equipped with hundreds of antenna elements. mmWave is poised to become a crucial component in the evolution of wireless networks [9]. However, it suffers from high path loss; thus, the use of beamforming to focus the signal on the intended user becomes a high-performance solution to overcome this issue [10].

Additionally, mMIMO technology employs multiple antennas at both the transmitter and receiver ends to significantly enhance communication effectiveness. By augmenting the number of antennas, mMIMO is capable of substantially improving signal strength, reducing interference, and increasing the capacity of wireless networks [11, 12].

*Corresponding author: m.oubassghir.ced@uca.ac.ma

The integration of these three technologies is crucial for advancing the next generation of wireless networks, opening up new possibilities in mobile communication [13, 14].

The paper is structured as follows: Section II discusses the system model employed. Section III presents the simulation results, and Section IV offers conclusions.

2 System model

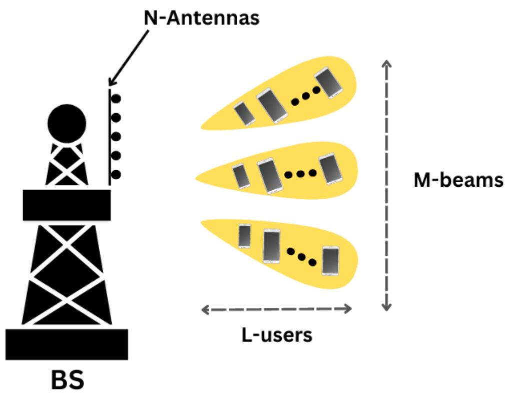

Figure 1. NOMA-mMIMO and mmWave System Model.

In this paper, we examine a system model depicted in Figure 1. This model features a single BS in the downlink process, equipped with N transmit antennas. The BS uses these antennas to generate M distinct beams in the spatial domain to serve L users ($L > M$), each equipped with a single antenna. The user set of the m-th group is denoted as U_m, where $U_i \cap U_j = \emptyset$ for $i \neq j$ and $\sum_{m=1}^{M} |U_m| = L$, with $|U_m|$ indicating the number of users in U_m.

The symbol transmitted to the l-th user within the m-th beam is represented by $x_{l,m}$ and is adjusted by the power coefficient $p_{l,m}$. Thus, by combining the signals from all beams, the downlink signal vector transmitted by the BS is effectively assembled:

$$s = \sum_{i=1}^{M} w_i \sum_{j=1}^{|U_i|} \sqrt{p_{j,i}} x_{j,i} \tag{1}$$

The beamforming vector for the m-th beam is denoted by w_m. The received signal from the l-th user in the m-th beam is created as follows:

$$y_{l,m} = \sum_{i=1}^{M} \sum_{j=1}^{|U_i|} w_i h_{j,i}^H \sqrt{p_{j,i}} x_{j,i} + v_{l,m} \tag{2}$$

In this framework, $h_{l,m} \in \mathbb{C}^{N\times1}$ describes the channel matrix that characterizes the link between the BS and the l-th user in the m-th beam. Meanwhile, $v_{l,m}$ represents the Additive White Gaussian Noise (AWGN) with a zero mean and a variance of σ^2.

To model the channel response vector $h_{l,m}$ of the l-th user in the m-th beam within mmWave communications, we utilize the widely adopted Saleh-Valenzuela model [15, 16]:

$$h_{l,m} = \sum_{c=1}^{C} g_{l,m}^{(c)} d(\theta_{l,m}^{(c)}) \tag{3}$$

where $g_{l,m}$ denotes the complex gain of the l-th user in the m-th beam, C represents the total number of multipath components, including both Line-of-Sight (LOS) and Non-Line-of-Sight (NLOS) components which result from reflection and scattering, and $\theta_{l,m}$ denotes the normalized Angle of Departure (AoD) for each beam, which is characterized as follows:

$$d(\theta_{l,m}) = [1, e^{-j\theta_{l,m}}, e^{-2j\theta_{l,m}}, \ldots, e^{-j(N-1)\theta_{l,m}}]^T \tag{4}$$

where:

$$\theta_{l,m} = 2\pi \frac{d}{\lambda} \sin(\phi_{l,m}) \tag{5}$$

λ represents the signal wavelength, d indicates the distance between antenna elements at BS, and $\phi_{l,m} \in [0, 2\pi]$ is the physical angle of departure.

To simplify our analysis, and without losing generality, we assume that:

$$||h_{1,m}^H w_m||_2 \geq ||h_{2,m}^H w_m||_2 \geq \ldots \geq ||h_{|U_m|,m}^H w_m||_2 \tag{6}$$

for $m = 1, 2, \ldots, M$ and $j = 1, 2, \ldots, |U_m|$. By adopting SIC in the m-th beam, the j-th user can detect and eliminate the i-th user's signals from its received signal for $1 \leq j < i < |U_m|$. Consequently, after SIC, the detailed signal received $y_{l,m}$ by the l-th user in the m-th beam can be expressed as:

$$\begin{aligned} y_{l,m} &= \sum_{i=1}^{M} \sum_{j=1}^{|U_i|} w_i h_{j,i}^H \sqrt{p_{j,i}} x_{j,i} + v_{l,m} \\ &= d_s + I_{Inter} + I_{Intra} + v_{l,m} \end{aligned} \tag{7}$$

where:

$$d_s = h_{l,m}^H w_m \sqrt{p_{l,m}} x_{l,m} \tag{8}$$

d_s represents the desired signal.

$$I_{Intra} = h_{l,m}^H w_m \sum_{j=1}^{l-1} \sqrt{p_{j,m}} x_{j,m} \tag{9}$$

I_{Intra} denotes intra-beam interference.

$$I_{Inter} = h_{l,m}^H \sum_{i\neq m} \sum_{j=1}^{|U_i|} w_i \sqrt{p_{j,i}} x_{j,i} \tag{10}$$

I_{Inter} indicates inter-beam interference and $v_{l,m}$ captures the noise effect.

2.1 SE and EE of NOMA-mMIMO and mmWave system

Based on the assumption (6) above and the expression in equation (7), the Signal-to-Interference-plus-Noise Ratio (SINR) for the l-th user in the m-th cluster is expressed as follows:

$$SINR_{l,m} = \frac{|h_{l,m}^H w_m|^2 p_{l,m}}{\gamma_{l,m}} \tag{11}$$

where:

$$\begin{aligned}\gamma_{l,m} &= h_{l,m}^H w_m \sum_{j=1}^{l-1} \sqrt{p_{j,m}} x_{j,m} \\ &+ h_{l,m}^H \sum_{i \neq m} \sum_{j=1}^{|U_i|} w_i \sqrt{p_{j,i}} x_{j,i} + \sigma^2\end{aligned} \tag{12}$$

For the l-th user in the m-th beam, the SE is expressed as follows:

$$SE_{l,m} = \log_2(1 + SINR_{l,m}) \tag{13}$$

The total SE is defined as:

$$SE = \sum_{m=1}^{M} \sum_{l=1}^{|U_m|} SE_{l,m} \tag{14}$$

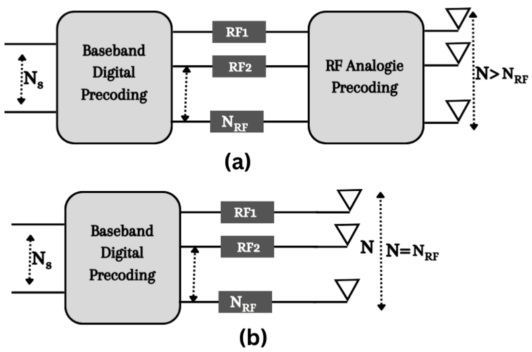

Figure 2. Beamforming Architecture: (a) Hybrid Structure. (b) Full digital structure.

Additionally, the EE performance, defined as the ratio of spectral efficiency to total power consumption, is described as follows:

$$EE_{l,m} = \frac{SE_{l,m}}{P + N_{RF} P_{RF} + N_{PS} P_{PS}} \tag{15}$$

The total EE is defined as:

$$EE = \sum_{m=1}^{M} \sum_{l=1}^{|U_m|} EE_{l,m} \tag{16}$$

P represents the transmission power, P_{RF} denotes the power consumed by each RF (Radio Frequency) chain, with N_{RF} and N_{PS} representing the total number of RF chains and phase shifters, respectively, and P_{PS} refers to the power consumption of each phase shifter.

Figure 2 represents both the full digital and hybrid beamforming structures. In the full digital structure, digital beamforming processes signals entirely in the digital domain, applying beamforming weights before converting them to analog for transmission. This method offers high flexibility and precise control but requires a separate RF chain for each antenna element, increasing complexity, cost, and power consumption.

Meanwhile, hybrid beamforming balances performance and complexity by combining digital and analog processing. Initial digital processing reduces the number of streams, which are then converted to analog. Further beamforming is done in the analog domain using phase shifters and combiners. This approach reduces the number of required RF chains, lowering cost and power consumption, but it offers less precise beam control and may involve performance trade-offs.

In the full digital structure, $N_{RF} = N$ and $N_{PS} = 0$, while in the hybrid structure, $N_{RF} = M$ and $N_{PS} = MN$ [14].

2.2 Beamforming Technique

To generate the beamforming vectors, employing Zero Forcing (ZF) can be a low-complexity approach as it effectively cancels the interference between users by applying the pseudo-inverse of the beamspace channel matrix.

However, ZF is contingent upon the condition that the number of users does not exceed the number of beams. In beamspace NOMA-mMIMO systems, where the number of users typically surpasses the number of beams, this becomes a limitation. To address this problem, Singular Value Decomposition (SVD) beamforming is proposed [15]. For a given matrix H_m that belongs to the complex space $\mathbb{C}^{N\times|U_m|}$, representing the beamspace channel matrix for all $|U_m|$ users in the m-th beam, i.e.,

$$H_m = [h_{1,m}, h_{2,m}, \ldots, h_{|U_m|,m}] \tag{17}$$

where $1 \leq m \leq M$. The matrix H_m can then be expressed as the product of these three matrices:

$$H_m = U_m S_m V_m^H \tag{18}$$

$U_m \in \mathbb{C}^{N\times N}$ and $V_m \in \mathbb{C}^{|U_m|\times|U_m|}$ are two complex matrices representing the left and right singular vectors of H_m, respectively. $S_m \in \mathbb{C}^{N\times|U_m|}$ is a diagonal matrix containing the singular values of the original channel matrix. w_m is chosen as the first column of U_m.

$$w_m = U_m^{(1)} \tag{19}$$

$U_m^{(1)}$ represents the initial column of U_m, denoting the left singular vector associated with the maximum singular value. In each beam, we select the user with the strongest connection to serve as the leader or head. This strategy leverages the head user's superior channel quality as a reference for generating the precoding vectors w_m, By using the head user's robust connection, the system can simplify and optimize the precoding process, ensuring that the generated vectors are highly efficient and effective. This approach not only streamlines the complex task of vector creation but also enhances overall system performance by minimizing interference and improving signal quality for all users within the beam.

3 Simulation results

In this section, we aim to evaluate the performance of NOMA-mMIMO and mmWave. Our analysis is conducted within a single cell, equipped with a BS that features $N = 64$ antennas configured in a ULA. The BS is designed to generate four beams ($M = 4$) and serve a total of $L = 32$ users. mmWave transmission is highly susceptible to obstructions due to its vulnerability to diffraction and significant path loss. Therefore, it primarily relies on LOS paths [17]. Consequently, we assign $C = 0$ for NLOS conditions, and consider $C = 1$ for LOS. For the power distribution, we selected: $P_{RF} = 250$ mW, $P_{PS} = 1$ mW, and $P = 1$ W [18]. The propagation loss for the channel model, with the carrier frequency set at 28 GHz, is described by the equation:
$PL_{l,m} = 61.4 + 20\log_{10}(D_{l,m})$ [19], where $D_{l,m}$ represent the distance between the BS and the l-th receiver in the m-th beam. The user distance depends on the cell size $S = 0.3$ Km, where $D_{l,m} = S\alpha_{l,m}$ with $\alpha_{l,m}$ being the factor that determines the distance between the user and the BS ($0 < \alpha_{l,m} \leq 1$). For generating the beamforming vectors w_m, we employed the SVD technique mentioned above, and for power allocation optimization we used the max-min fairness method described in [20], which ensures equitable rate distribution among users by maximizing the minimum achievable throughput. In this paper, we consider that the signal-to-noise ratio (SNR) is given by $\frac{P}{\sigma^2}$. The simulation results are obtained using MATLAB 2022.

To compare the achieved throughput, we will consider an FDMA-mMIMO and mmWave system. Users will be assigned to M groups, with each group performing frequency division multiple access. Subsequently, the spectral efficiency of the FDMA-mMIMO-mmWave scheme for the l-th user is [16]:

$$SE_l^{FDMA} = \frac{1}{|U^l|}\log_2\left(1 + \frac{|h_l^H w_l|^2 p_l}{\sum_{j\neq|U_l|}|h_j^H w_j|^2 p_j + \frac{\sigma^2}{|U^l|}}\right) \quad (20)$$

where U^l denotes the group to which the l-th user belongs, and $|U^l|$ represents the number of users in this group. For the power allocation, we used a fixed power allocation scheme, distributing the available power equally among the users, and the beamforming vector was generated based on the SVD beamforming method discussed above.

Figure 3 provides an insightful overview of the spectral efficiency trends for NOMA and FDMA, both integrated with mMIMO and mmWave, as SNR increases. The simulation results demonstrate a marked improvement in spectral efficiency when implementing NOMA compared to FDMA. This efficiency boost showcases the capacity of NOMA to handle high user density without compromising the quality of service, thereby enhancing network throughput and user experience.

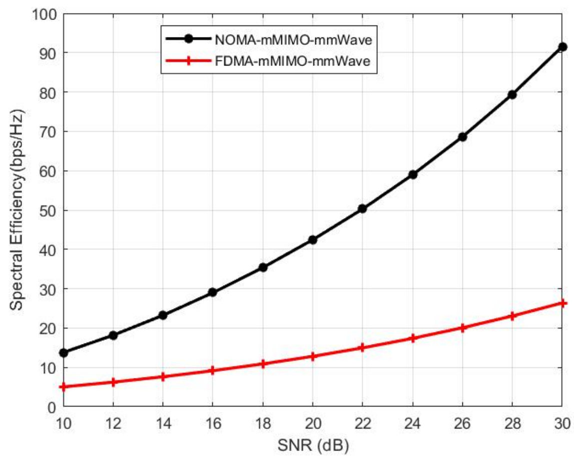

Figure 3. Comparison of Spectral Efficiency: NOMA-mMIMO-mmWave vs. FDMA-mMIMO-mmWave, using Hybrid Architecture.

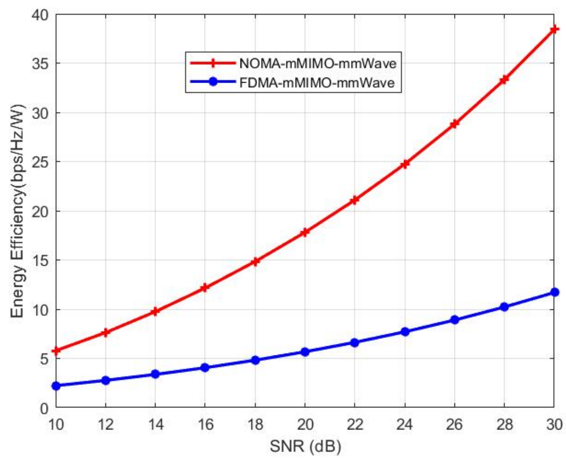

Figure 4. Energy Efficiency Comparison of NOMA-mMIMO-mmWave and FDMA-mMIMO-mmWave, using Hybrid Architecture.

Furthermore, Figure 4 explores the energy efficiency aspect, presenting a clear advantage of NOMA over FDMA. By highlighting the ability of NOMA to achieve higher energy efficiency, this figure emphasizes the technology's effectiveness in reducing power consumption while maintaining or enhancing data transmission capabilities.

Thanks to the hybrid structure, which balances energy consumption and throughput, the effectiveness of NOMA in serving a high number of users with a low number of RF chains is significantly boosted. Figure 5 demonstrates this by showing that, for high spectral efficiency values, the hybrid structure makes a noticeable difference in energy consumption compared to a full digital structure. In a full digital setup, the number of RF chains required is equal to the number of antennas, which are major contributors to energy consumption. In a mMIMO system, this num-

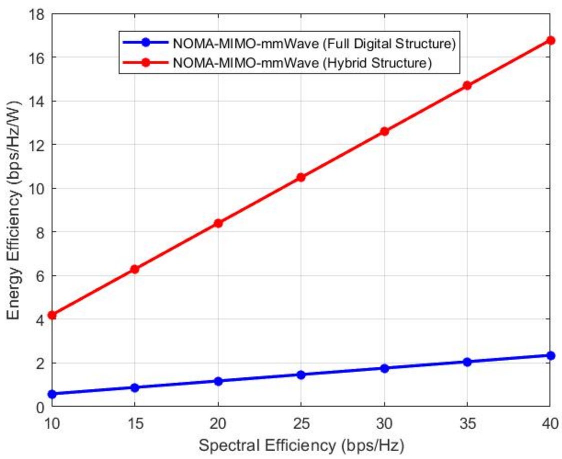

Figure 5. Performance Comparison of NOMA-mMIMO and mmWave Using Hybrid and Full Digital Architectures.

ber can be on the order of hundreds, leading to substantial energy consumption.

The collective findings from these figures emphasize the superior performance metrics of NOMA, not only in terms of spectral and energy efficiencies but also in its capability to revolutionize telecommunications infrastructure. NOMA stands out as a pivotal technology for 5G deployments and future telecommunications standards, particularly in densely populated urban areas and scenarios requiring massive device connectivity.

4 Conclusion

In this study, we investigated the achieved spectral and energy efficiency of NOMA-mMIMO with mmWave technology, specifically focusing on the implementation of beamforming techniques. The analysis demonstrated that the integration of NOMA with mMIMO and mmWave technologies can significantly bolster both spectral and energy efficiencies. The use of mmWave frequencies provides ample bandwidth, while mMIMO technology enhances signal quality and system capacity through the use of multiple antennas. When combined with NOMA, these technologies facilitate the accommodation of a higher number of users and the provision of higher data rates without the need for additional spectrum resources.

References

[1] W. Jiang, B. Han, M. A. Habibi and H. D. Schotten, “The Road Towards 6G: A Comprehensive Survey," in IEEE Open Journal of the Communications Society, **2**, pp. 334-366, (2021).

[2] V. K. Quy, A. Chehri, N. M. Quy, N. D. Han and N. T. Ban, “Innovative Trends in the 6G Era: A Comprehensive Survey of Architecture, Applications, Technologies, and Challenges". IEEE Access, **11**, pp. 39824-39844, (2023).

[3] M. A. Gupta and R. K. Jha, “A Survey of 5G Network: Architecture and Emerging Technologies," in IEEE Access, **3**, pp. 1206-1232, (2015).

[4] M.Oubassghir and M.Boulouird. “Capacity Analysis of NOMA with mmWave for 5G wireless Communications and Beyond". Proceedings of the 6th International Conference on Networking, Intelligent Systems and Security (NISS). Association for Computing Machinery, New York, NY, USA, Article 43, pp. 1–5, (2023).

[5] M. Moltafet, N. M. Yamchi, M. R. Javan and P. Azmi, “Comparison Study Between PD-NOMA and SCMA". IEEE Transactions on Vehicular Technology, **67**, no. 2, pp. 1830-1834, (Feb 2018).

[6] Umar Ghafoor, Mudassar Ali, Humayun Zubair Khan, Adil Masood Siddiqui, Muhammad Naeem, “NOMA and future 5G and B5G wireless networks". Journal of Network and Computer Applications, **204**, (2022).

[7] B. Makki, K. Chitti, A. Behravan and M. -S. Alouini, “A Survey of NOMA: Current Status and Open Research Challenges". IEEE Open Journal of the Communications Society, **1**, pp. 179-189, (2020).

[8] A. N. Uwaechia and N. M. Mahyuddin, “A Comprehensive Survey on Millimeter Wave Communications for Fifth-Generation Wireless Networks: Feasibility and Challenges". IEEE Access, **8**, pp. 62367-62414, (2020).

[9] L. Zhu, Z. Xiao, X. -G. Xia and D. Oliver Wu, “Millimeter-Wave Communications With Non-Orthogonal Multiple Access for B5G/6G". IEEE Access, **7**, pp. 116123-116132, (2019).

[10] Z. Xiao et al., “A Survey on Millimeter-Wave Beamforming Enabled UAV Communications and Networking". IEEE Communications Surveys and Tutorials, **24**, no. 1, pp. 557-610, (2022).

[11] J. Zhang, E. Björnson, M. Matthaiou, D. W. K. Ng, H. Yang and D. J. Love, “Prospective Multiple Antenna Technologies for Beyond 5G". IEEE Journal on Selected Areas in Communications, **38**, no. 8, pp. 1637-1660, (Aug 2020).

[12] A. Aalili and M. Boulouird, “Spectral and Energy Efficiency in Cell-free Massive MIMO systems for 5G and 6G Wireless Communications," 2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), Tenerife, Canary Islands, Spain, pp. 1-6, (2023).

[13] J. Ghosh, V. Sharma, H. Haci, S. Singh and I. -H. Ra, “Performance Investigation of NOMA Versus OMA Techniques for mmWave Massive MIMO Communications". IEEE Access, **9**, pp. 125300-125308, (2021).

[14] L. Pang et al., “Joint Power Allocation and Hybrid Beamforming for Downlink mmWave-NOMA Systems". IEEE Transactions on Vehicular Technology, **70**, no. 10, pp. 10173-10184, (Oct 2021).

[15] B. Wang, L. Dai, Z. Wang, N. Ge, and S. Zhou, “Spectrum and energy efficient beamspace MIMO-

NOMA for millimeterwave communications using lens antenna array”. IEEE Journal on Selected Areas in Communications, **35**, no. 10, pp.2370–2382, (Oct 2017).

[16] L. Dai, B. Wang, M. Peng, and S. Chen, “Hybrid precoding-based millimeter-wave massive MIMO-NOMA with simultaneous wireless information and power transfer”. IEEE Journal on Selected Areas in Communications, **37**,no. 1, pp. 131–141, (Jan. 2019).

[17] D. Zhang, Z. Zhou, C. Xu, Y. Zhang, J. Rodriguez and T. Sato, “Capacity Analysis of NOMA With mmWave Massive MIMO Systems,” in IEEE Journal on Selected Areas in Communications, **35**, no. 7, pp. 1606-1618, (July 2017).

[18] X. Gao, L. Dai, S. Han, C. I, and R. W. Heath, “Energy-efficient hybrid analog and digital precoding for mmwave MIMO systems with large antenna arrays,” IEEE Journal on Selected Areas in Communications, **34**, no. 4, pp. 998–1009, (April 2016).

[19] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar and L. Hanzo, “Millimeter-Wave Communications: Physical Channel Models, Design Considerations, Antenna Constructions, and Link-Budget," in IEEE Communications Surveys and Tutorials, **20**, no. 2, pp. 870-913, (2018).

[20] R. Jiao, L. Dai, W. Wang, F. Lyu, N. Cheng and X. Shen, “Max-Min Fairness for Beamspace MIMO-NOMA: From Single-Beam to Multi-Beam," in IEEE Transactions on Wireless Communications, **21**, no. 2, pp. 739-752, (Feb. 2022).