Abstract—In this paper, we compare the individual rate of MIMO-NOMA and MIMO-OMA when users are paired into clusters. A power allocation (PA) strategy is proposed, which ensures that MIMO-NOMA achieves a higher individual rate for each user than MIMO-OMA with arbitrary PA and optimal degrees of freedom split. In addition, a special case with equal degrees of freedom and arbitrary PA for OMA is considered, for which the individual rate superiority of NOMA still holds. Moreover, it is shown that NOMA can attain better fairness through appropriate PA. Finally, simulations are carried out, which validate the developed analytical results.

I. INTRODUCTION

The non-orthogonal multiple access (NOMA) has drawn great attention as a promising technology for improving the spectral efficiency for the next generation mobile communication networks [1]-[6]. There exist two main NOMA schemes, i.e., power-domain and code-domain NOMA. In this paper, we focus on the former, in which users are multiplexed on power domain. For notational simplicity, we refer to power-domain NOMA as NOMA.

A few works have verified via simulation the superiority of NOMA over OMA for multi-user scenarios in term of achievable sum rate [7]-[10]. For single-input single-output (SISO) systems, [7] shows that NOMA can achieve a larger sum rate, while [8] illustrates that a larger ergodic sum rate is obtained by NOMA for a cellular downlink system with randomly deployed users. As for multiple-input multiple-output (MIMO) systems, [9], [10] provide some insight: [9] verifies that a larger ergodic sum rate for two users can be obtained by NOMA, whereas [10] shows that NOMA can achieve a larger sum rate for a multi-user scenario, with two users paired into a cluster, and sharing a common transmit beamforming vector.

Some recent works aim to analytically prove that NOMA achieves higher sum rate over OMA. For SISO systems, power allocation (PA) in [11] is conducted to guarantee that NOMA achieves a larger sum rate than OMA with equal power coefficients and degrees of freedom (DoF). For MIMO systems, [12] derives the sum rate gain of NOMA over OMA under two extreme cases of user pairing: 1) the best user with the worst user; 2) the best user with the second best user. Moreover, a cognitive radio inspired PA is proposed, which ensures that the data rate of the weak user is larger than that in OMA. However, the sum rate for OMA is not optimized in the above works as equal power and DoF are allocated to users.

In [13], [14], the authors overcome this issue, and demonstrate that NOMA achieves a larger sum rate than OMA for scenarios with two users and multiple users per cluster, respectively.

The major drawback of the sum rate comparison is that it neglects fairness. To the best of our knowledge, none of the previous works considers fairness during sum rate comparison. Note that although simulation results in [13] show that NOMA achieves higher fairness, no theoretical analysis is provided. Hence, in order to account for fairness, we need to compare the individual rates of the users. In particular, the individual rate for any user in NOMA should be higher or equal than its counterpart in OMA. In [15], the PA scheme for a SISO system is designed such that the individual rate of each user in NOMA is guaranteed to be larger than its counterpart in OMA. However, [15] still adopts equal PA and DoF for OMA, which is suboptimal. By filling in this gap, the main contribution of this paper lies in:

- A general and fair individual rate comparison is considered, in which the PA for OMA is arbitrary and the DoF is split such that the maximum sum rate in OMA is achieved. On this basis, a PA strategy is proposed, which ensures that NOMA achieves higher individual rates than OMA.
- For the particular case with equal DoF and arbitrary PA, analytical results are provided to demonstrate the superiority of NOMA over OMA in terms of individual rates.
- In addition to the individual rate superiority, it is also shown that better fairness is achieved by NOMA through appropriate PA.

The rest of the paper is organized as follows. The system model is introduced in Section II. The individual rate comparison between MIMO-NOMA and MIMO-OMA is conducted in Section III, where a PA strategy is additionally proposed. The particular case of equal DoF is also discussed in Section III, while simulation results are shown in Section IV. Conclusions are finally drawn in Section V.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

A multi-user MIMO-NOMA downlink transmission scenario is investigated, in which a micro base station (BS) deployed with $M$ antennas sends information to $2M$ users,
each with $N$ antennas. Two users are paired into a cluster for complexity reduction \([12]\), and NOMA is only applied between them. According to this, there are $M$ clusters in the system. We adopt the block fading channel model, where both path loss component and small scale fading are considered, e.g., the channel matrix from the BS to user $k,k \in \{1,2\}$ in cluster $m,m \in \{1,\ldots,M\}$ is $H_{m,k} = G_{m,k}/L_{m,k}$, with $G_{m,k} \in \mathbb{C}^{N \times M}$ denoting the Rayleigh fading channel matrix and $L_{m,k}$ representing the path loss component. The transmit and receive beamforming vectors fulfill the following conditions \([10]\): 1) zero-forcing (ZF) precoding is conducted at the BS to remove the inter-cluster interference; 2) signal alignment is conducted at the receiver between users in the same cluster, i.e., $v_{m,2}^H G_{m,2} = v_{m,1}^H G_{m,1}$, where $v_{m,k}^H$ denotes the receive beamforming vector.

As users in the same cluster share a common transmit beamforming vector, the signal transmitted from the BS can be expressed as
\[
x = Ps,
\]
where $P = [p_1 \cdots p_M] \in \mathbb{C}^{M \times M}$, with $p_m \in \mathbb{C}^{M \times 1}$ representing the normalized transmit beamforming vector for cluster $m$. Additionally, the information bearing vector $s \in \mathbb{C}^{M \times 1}$ is given by
\[
s = \begin{bmatrix} \alpha_{1,1}s_{1,1} + \alpha_{1,2}s_{1,2} \\ \vdots \\ \alpha_{M,1}s_{M,1} + \alpha_{M,2}s_{M,2} \end{bmatrix},
\]
where $s_{m,k}$ and $\alpha_{m,k}$ represent the signal and corresponding PA coefficient for user $(m,k)$, respectively, satisfying $\alpha_{m,1}^2 + \alpha_{m,2}^2 = 1, \forall m$.

At the receiver of user $(m,k)$, the normalized receive beamforming vector $v_{m,k}$ is applied, and thus, the received signal $y_{m,k}$ is given by
\[
y_{m,k} = v_{m,k}^H y_{m,k} = \alpha_{m,1}v_{m,1}^H H_{m,1} p_m s_{m,1} + \alpha_{m,2} v_{m,2}^H H_{m,2} p_{m} s_{m,2} + \sum_{\ell \neq m}^M \alpha_{m,\ell} v_{m,\ell}^H H_{m,\ell} p_{m} s_{m,\ell} + v_{m,k}^H n_{m,k},
\]
where $(\cdot)^H$ represents the Hermitian transpose operation and $n_{m,k} \in \mathbb{C}^{N \times 1}$ follows $\mathcal{CN}(0,\sigma^2_0 I)$ as the additive white Gaussian noise (AWGN) vector at user $(m,k)$.

As ZF precoding is adopted at the BS, inter-cluster interference can be eliminated, and thus, the cluster index $m$ can be dropped for notational simplicity. Consequently, the received signal can be rewritten as
\[
y_k = y_{1,k} + y_{2,k} = \alpha_{1,k} v_{1,k}^H H_{1,1} p_{1,k} s_{1,k} + \alpha_{2,k} v_{2,k}^H H_{2,2} p_{2,k} s_{2,k} + v_{1,k}^H n_{1,k} + v_{2,k}^H n_{2,k}.
\]

Without loss of generality, the effective channel gains of the users are ordered as follows:
\[
|v_{1,k}^H H_{1,1} p_{1,k}|^2 \geq |v_{2,k}^H H_{2,2} p_{2,k}|^2.
\]

Accordingly, successive interference cancellation (SIC) is conducted at user 1 to remove the interference from user 2, and because of this, the achieved data rate at user 1 can be expressed as \([13]\)
\[
R_{1}^{\text{NOMA}} = \log_2 \left( 1 + \rho \frac{|v_{1,k}^H H_{1,1} p_{1,k}|^2}{E[\rho |n_{1,k}|^2]} \right),
\]
where $\rho = \frac{1}{E[|v_{2,k}^H n_{2,k}|^2]}$ is the same for the two users, as the receive beamforming vector is normalized and the noise variance remains unchanged after rotation. $E[\cdot]$ denotes the expectation operator.

In contrast, user 2 considers user 1’s signal as interference, and thus, its achievable rate is given by
\[
R_{2}^{\text{OMA}} = \log_2 \left( 1 + \frac{\rho_1 |v_{2,k}^H H_{2,2} p_{2,k}|^2}{\rho_2 |v_{2,k}^H H_{2,2} p_{1,k}|^2} \right),
\]
where $\rho_1$ and $\rho_2$, satisfying $\alpha_1^2 + \alpha_2^2 = 1$, the split of the DoF between the two users is optimized to achieve the maximum sum rate for fair comparison. We use $\lambda_1$ and $\lambda_2$ to denote the fractions of the DoF for users 1 and 2, respectively, which should satisfy $\lambda_1 + \lambda_2 = 1$. As such, the achievable rate at user $k$ can be expressed as \([13]\)
\[
R_{k}^{\text{OMA}} = \lambda_k \log_2 \left( 1 + \frac{\rho_2 |v_{2,k}^H H_{2,2} p_{k}|^2}{\sum_{l=1}^{2} \alpha_l^2 |v_{l,k}^H H_{l,1} p_{k}|^2} \right).
\]

Now, the sum rate for the two users in the same cluster in MIMO-OMA is given by \([13]\) Lemma 1
\[
R_{\text{sum}}^{\text{OMA}} \leq \log_2 \left( 1 + \frac{\sum_{k=1}^{2} \rho_2 |v_{2,k}^H H_{2,2} p_{k}|^2}{\sum_{l=1}^{2} \alpha_l^2 |v_{l,k}^H H_{l,1} p_{k}|^2} \right).
\]

Note that when \([10]\) is satisfied, the maximum sum rate for OMA is achieved, and the corresponding individual rates for users 1 and 2 are used for OMA to ensure a fair comparison.

**B. Problem Formulation**

In \([13]\), the authors prove that NOMA can achieve a larger sum rate than OMA by simply assigning the same power coefficients to both schemes. However, having a higher sum rate does not guarantee that each user in NOMA has a higher data rate than its counterpart in OMA. Indeed, it is easy to come up with an instance in which the data rate of the weak user (user 2) in NOMA is below its counterpart in OMA if simply assigning the same power coefficients. For example, if $\rho_1 |v_{1,k}^H H_{1,1} p_{1,k}|^2 = \rho_2 |v_{2,k}^H H_{2,2} p_{2,k}|^2 = 0.25$ and $\alpha_1^2 = 0.5 \alpha_2^2$, then
\[
\log_2 \left( 1 + \frac{\rho_2 |v_{2,k}^H H_{2,2} p_{2,k}|^2}{1 + \rho_2 |v_{2,k}^H H_{2,2} p_{1,k}|^2} \right) = \log_2(1.22) < \log_2(1.23) = \frac{\alpha_1^2 |v_{1,k}^H H_{1,1} p_{1,k}|^2}{\alpha_1^2 |v_{1,k}^H H_{1,1} p_{1,k}|^2 + \rho_2 |v_{2,k}^H H_{2,2} p_{2,k}|^2} \log_2 \left( 1 + \frac{\rho_2 |v_{2,k}^H H_{2,2} p_{2,k}|^2}{\rho_2 |v_{2,k}^H H_{2,2} p_{1,k}|^2} \right),
\]
which means that NOMA may lead to unfair data rate between its two users when compared with OMA. Consequently, to further verify the superiority of NOMA over OMA, PA should be conducted such that the data rate of each user in NOMA exceeds its counterpart in OMA. A
PA scheme satisfying this requirement is proposed in [15]. However, [15] adopts time-division multiple access with equal power and DoF for its users as the representative of OMA, which does not achieve maximum sum rate for OMA. On the other hand, for a general case, like any PA for OMA, does this conclusion still hold? To the best of our knowledge, this problem has never been considered in the literature.

To validate that NOMA achieves a higher individual rate than OMA for an arbitrary PA in OMA, we need to find the feasible power coefficients for NOMA, which achieve this goal under any given power coefficients and optimal DoF for OMA. The considered problem can be formulated as follows:

\[
\begin{align*}
\text{find } & \quad \alpha_1, \alpha_2 \\
\text{s.t. } & \quad (10), \\
& \quad \log_2 \left( 1 + \rho \alpha_1^2 |v_1^H H_1 p|^2 \right) \\
& \quad \quad \geq \lambda_1 \log_2 \left( 1 + \frac{\rho \alpha_2^2 |v_1^H H_1 p|^2}{|v_1^H H_1 p|^2} \right), \\
& \quad \log_2 \left( 1 + \frac{\rho \alpha_2^2 |v_2^H H_2 p|^2}{1 + \rho \alpha_2^2 |v_2^H H_2 p|^2} \right) \\
& \quad \quad \geq \lambda_2 \log_2 \left( 1 + \frac{\rho \alpha_2^2 |v_2^H H_2 p|^2}{1 + \rho \alpha_2^2 |v_2^H H_2 p|^2} \right), \\
& \quad \alpha_1^2 + \alpha_2^2 = 1, \alpha_1^2, \alpha_2^2 \in [0, 1],
\end{align*}
\]

where (11b) ensures that OMA achieves the maximum sum rate, while (11c) and (11d) guarantee that NOMA outperforms OMA for both users.

III. PROPOSED PA SCHEME

A. Optimal DoF and Varying Power

In this section, we propose a PA strategy, which satisfies the constraints (11b)-(11d). First, with some algebraic manipulations on (11b) and (11d), the PA strategy for NOMA is given by

\[
\begin{align*}
\alpha_1^2 & \geq \frac{(1 + \rho \alpha_1^2 |v_1^H H_1 p|^2)^{\lambda_1} - 1}{\rho |v_1^H H_1 p|^2}, \\
\alpha_2^2 & \leq \frac{1 + \rho |v_2^H H_2 p|^2 - (1 + \frac{\rho \alpha_2^2 |v_2^H H_2 p|^2}{\lambda_2})^{\lambda_2}}{\rho |v_2^H H_2 p|^2 (1 + \frac{\rho \alpha_2^2 |v_2^H H_2 p|^2}{\lambda_2})^{\lambda_2}}.
\end{align*}
\]

Now, to ensure a feasible solution for \( \alpha_1^2 \), the following condition must be satisfied:

\[
\begin{align*}
(1 + \frac{\rho \alpha_1^2 |v_1^H H_1 p|^2}{\lambda_1} - 1) & \leq \frac{1 + \rho |v_2^H H_2 p|^2 - (1 + \frac{\rho \alpha_2^2 |v_2^H H_2 p|^2}{\lambda_2})^{\lambda_2}}{\rho |v_2^H H_2 p|^2 (1 + \frac{\rho \alpha_2^2 |v_2^H H_2 p|^2}{\lambda_2})^{\lambda_2}}.
\end{align*}
\]

With the help of (11b), and after some algebraic manipulations, (13) can be further expressed as

\[
\begin{align*}
(1 + \frac{\rho \alpha_1^2 |v_1^H H_1 p|^2}{\lambda_1} - 1) & \leq \frac{1 + \rho |v_2^H H_2 p|^2 - (1 + \frac{\rho \alpha_2^2 |v_2^H H_2 p|^2}{\lambda_2})^{\lambda_2}}{\rho |v_2^H H_2 p|^2 (1 + \frac{\rho \alpha_2^2 |v_2^H H_2 p|^2}{\lambda_2})^{\lambda_2}}.
\end{align*}
\]

B. Equal DoF and Varying Power

In (11), the DoF are split according to (10). As the PA is arbitrary, the resulting fractions of DoF can also take any value, which may be infeasible to realize in practice [16]. Motivated by this observation, in this section, we consider a simple and practical case when the DoF for two users in the same cluster in MIMO-OMA are equal, while the PA is still arbitrary. Compared with [15], the considered case is more general as the PA can be arbitrary. In contrast to [12], which only ensures the QoS of the weak user, the considered case takes into account both strong and weak users.

The corresponding problem can be formulated as:

\[
\begin{align*}
\text{find } & \quad \alpha_1, \alpha_2 \\
\text{s.t. } & \quad \log_2 \left( 1 + \rho \alpha_1^2 |v_1^H H_1 p|^2 \right) \\
& \quad \quad \geq 1/2 \log_2 \left( 1 + 2 \rho \alpha_1^2 |v_1^H H_1 p|^2 \right), \\
& \quad \log_2 \left( 1 + \frac{\rho \alpha_2^2 |v_2^H H_2 p|^2}{1 + \rho \alpha_2^2 |v_2^H H_2 p|^2} \right) \\
& \quad \quad \geq \frac{1}{2} \log_2 \left( 1 + 2 \rho \alpha_2^2 |v_2^H H_2 p|^2 \right), \\
& \quad \alpha_1^2 + \alpha_2^2 = 1, \alpha_1^2, \alpha_2^2 \in [0, 1].
\end{align*}
\]

Note that the main difference between (15) and (11) lies in the fact that (11b) is no longer a constraint in the former. Instead, both \( \lambda_1 \) and \( \lambda_2 \) take a fixed value of \( \frac{1}{2} \).
To find the solution of (15), we start with the case when equality is attained in (15b). Accordingly, we have
\[
\frac{(1 + \rho |v_2^H H_2 p|^2)^2}{(1 + \rho \alpha_1^2 |v_2^H H_2 p|^2)^2} = 1 + 2\rho(1 - \alpha_1^2)|v_2^H H_2 p|^2 \tag{16a}
\]
\[
\iff \alpha_1^2 = \frac{1 + \rho |v_2^H H_2 p|^2 - \sqrt{1 + 2\rho \alpha_2^2 |v_2^H H_2 p|^2}}{\rho |v_2^H H_2 p|^2 + \sqrt{1 + 2\rho \alpha_2^2 |v_2^H H_2 p|^2}}. \tag{16b}
\]

On this basis, we ensure that (15b) always holds. To achieve that, we rewrite (15b) as
\[
(1 + \rho \alpha_1^2 |v_1^H H_1 p|^2)^2 \geq 1 + 2\rho \alpha_1^2 |v_1^H H_1 p|^2, \tag{17}
\]
and (16b) as
\[
1 + 2\rho \alpha_1^2 |v_2^H H_2 p|^2 = 2(1 + \rho |v_2^H H_2 p|^2) - (1 + \rho |v_2^H H_2 p|^2)^2
\]
\[
\leq 2(1 + \rho |v_2^H H_2 p|^2)(1 + \rho \alpha_1^2 |v_1^H H_1 p|^2)
\]
\[
\leq 2(1 + \rho |v_2^H H_2 p|^2) - 2(1 + \rho |v_2^H H_2 p|^2)
\]
\[
= (1 + \rho \alpha_1^2 |v_2^H H_2 p|^2)^2, \tag{18}
\]
where the inequality comes from the Jensen’s inequality.

Now, with the help of (15) and (18), we obtain
\[
(1 + \rho \alpha_1^2 |v_1^H H_1 p|^2)^2 - 1 - 2\rho \alpha_1^2 |v_1^H H_1 p|^2
\]
\[
\geq (1 + \rho \alpha_1^2 |v_2^H H_2 p|^2)^2 - 1 - 2\rho \alpha_1^2 |v_2^H H_2 p|^2 \tag{19}
\]
\[
\geq 0,
\]
which is exactly (17). Hence, (15b) always holds.

Similarly, we can prove that when equality is achieved for (15b), (15a) holds. In this case, we have the PA strategy for NOMA as
\[
\alpha_1^2 = \frac{\sqrt{1 + 2\rho \alpha_2^2 |v_1^H H_1 p|^2} - 1}{\rho |v_1^H H_1 p|^2}. \tag{20}
\]

Clearly, when \(\alpha_1^2\) lies in the boundary between the values in (16) and (20), MIMO-NOMA always achieves higher individual rates than MIMO-OMA.

IV. SIMULATION RESULTS

In this section, simulations are conducted to compare the individual rates of MIMO-NOMA and MIMO-OMA, and hence, verify the accuracy of the developed analytical results. In simulations, \(M = 4\) and the path-loss exponent is 3.8.

Fig. 1 compares the individual rate between MIMO-NOMA and MIMO-OMA with equal DoF, when the power coefficient for the weak user varies. In simulations, \(\rho = 30\) dB, \(|v_1^H H_1 p|^2 = 0.052\) and \(|v_2^H H_2 p|^2 = 0.0052\). Note that NOMA_1 and NOMA_2 denote the cases when the power coefficient of the strong user in MIMO-NOMA satisfies (20) and (16b), respectively. As expected, \(R_1\) in NOMA_1 equals that in OMA, while \(R_2\) in NOMA_2 is the same as that in OMA.

Moreover, both \(R_2\) in NOMA_1 and \(R_1\) in NOMA_2 exceed their counterparts in OMA, which verifies the superiority of NOMA over OMA in terms of individual rate. Particularly, when \(\alpha_2^2 \in [0, 0.8]\), it can be seen that \(R_1^{\text{NOMA}_1} = R_1^{\text{OMA}} > R_2^{\text{NOMA}_1} > R_2^{\text{OMA}}\). Therefore, NOMA_1 also provides better fairness than OMA.

Figs. 2-4 present results obtained when the optimal DoF is used for OMA. The legends NOMA_1 and NOMA_4 denote the scenarios when \(\alpha_1^2\) follows (12b) and (12d), respectively. In addition, the legends OMA_1 and OMA denote the OMA scheme in (15) (with equal power and DoF) and the one considered in this paper (with arbitrary power and optimal DoF), respectively.

In Fig. 2, we show how \(R_1\) varies with \(\rho\) for the previously mentioned four schemes. It can be seen that NOMA_4 achieves the highest rate for \(R_1\), while OMA in (15) obtains the lowest rate. In addition, NOMA_4 attains the same rate as OMA.
Morover, NOMA also outperforms OMA \cite{15}, as both fairness is achieved by NOMA when compared with OMA. Particularly, under NOMA, we can easily conclude that NOMA can always achieve higher individual rates than OMA considered in this paper, once (12) is satisfied. Particularly, under NOMA, better fairness is achieved by NOMA when compared with OMA. Moreover, NOMA also outperforms OMA \cite{15}, as both $R_1$ and $R_2$ for NOMA are higher than that for OMA \cite{15}

Lastly, from Fig. 4, we can observe that OMA considered in this paper has a larger sum rate than OMA \cite{15} owing to the use of optimal DoF. This justifies the necessity of optimizing the DoF for the comparison between NOMA and OMA. The order of the sum rate is $\text{NOMA}_3 > \text{NOMA}_4 > \text{OMA} > \text{OMA}$ \cite{15}. NOMA$_3 >$ NOMA$_4$ can be explained by the fact that allocating more power to the stronger user results in a higher sum rate.

V. CONCLUSION

A fair individual rate comparison between MIMO-NOMA and MIMO-OMA has been investigated. We have proposed a PA strategy, which guarantees that MIMO-NOMA achieves a higher individual rate than MIMO-OMA with arbitrary power coefficients and optimized DoF. Additionally, we have shown that this also holds for the case of equal DoF and arbitrary power coefficients. Numerical results verify the accuracy of the developed analytical results.

REFERENCES

[1] S. M. R. Islam, M. Zeng, and O. A. Dobre, “NOMA in 5G systems: Exciting possibilities for enhancing spectral efficiency,” IEEE 5G Tech. Focus, vol. 1, no. 2, May 2017. [Online]. Available: 307 http://5g.ieee.org/tech-focus.

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

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

[4] 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.

[5] M. Zeng, G. I. Tsiropoulos, O. A. Dobre, and M. H. Ahmed, “Power allocation for cognitive radio networks employing non-orthogonal multiple access,” in Proc. IEEE Global Telecommun. Conf., Washington DC, USA, Dec. 2016, pp. 1–5.

[6] M. Zeng, G. I. Tsiropoulos, A. Yadav, O. A. Dobre, and M. H. Ahmed, “A two-phase power allocation scheme for CRNs employing NOMA,” in Proc. IEEE Global Telecommun. Conf., to appear.

[7] Y. Saito, et al., “Non-orthogonal multiple access (NOMA) for cellular future radio access,” in Proc. IEEE Veh. Technol. Conf., Dresden, Germany, Jun. 2013, pp. 1–5.

[8] 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.

[9] Q. Sun, S. Han, I. Chih-Lin, and Z. Pan, “On the ergodic capacity of MIMO NOMA systems,” IEEE Wireless Commun. Lett., vol. 4, no. 4, pp. 405–408, Dec. 2015.

[10] Z. Ding, R. Schober, and H. V. Poor, “A general MIMO framework for NOMA downlink and uplink transmission based on signal alignment,” IEEE Trans. Wireless Commun., vol. 15, no. 6, pp. 4438–4454, Jun. 2016.

[11] M. M. El-Sayed, A. S. Ibrahim, and M. M. Khairy, “Power allocation strategies for non-orthogonal multiple access,” in Proc. MoWNet, Cairo, Egypt, Apr. 2016, pp. 1–6.

[12] Z. Ding, F. Adachi, and H. V. Poor, “The application of MIMO to non-orthogonal multiple access,” IEEE Trans. Wireless Commun., vol. 15, no. 1, pp. 537–552, Jan. 2016.

[13] M. Zeng, A. Yadav, O. A. Dobre, G. I. Tsiropoulos, and H. V. Poor, “On the sum rate of MIMO-NOMA and MIMO-OMA systems,” IEEE Wireless Commun. Lett., vol. PP, no. 99, pp. 1–1, Jun. 2017.

[14] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, and H. V. Poor, “Capacity comparison between MIMO-NOMA and MIMO-OMA with multiple users in a cluster,” IEEE J. Sel. Areas Commun., DOI: 10.1109/JSAC.2017.2725879, Jul. 2017.

[15] Z. Yang, Z. Ding, P. Fan, and N. Al-Dhahir, “A general power allocation scheme to guarantee quality of service in downlink and uplink NOMA systems,” IEEE Trans. Wireless Commun., vol. 15, no. 11, pp. 7244–7257, Nov. 2016.

[16] S. Shi, L. Yang, and H. Zhu, “Outage balancing in downlink nonorthogonal multiple access with statistical channel state information,” IEEE Trans. Wireless Commun., vol. 15, no. 7, pp. 4718–4731, Jul. 2016.