NOMA: An Information Theoretic Perspective

Peng Xu, Zhiguo Ding, *Member, IEEE*, Xuchu Dai and H. Vincent Poor, *Fellow, IEEE*

Abstract—In this letter, the performance of non-orthogonal multiple access (NOMA) is investigated from an information theoretic perspective. The relationships among the capacity region of broadcast channels and two rate regions achieved by NOMA and time-division multiple access (TDMA) are illustrated first. Then, the performance of NOMA is evaluated by considering TDMA as the benchmark, where both the sum rate and the individual user rates are used as the criteria. In a wireless downlink scenario with user pairing, the developed analytical results show that NOMA can outperform TDMA not only for the sum rate but also for each user's individual rate, particularly when the difference between the users' channels is large.

I. INTRODUCTION

Because of its superior spectral efficiency, non-orthogonal multiple access (NOMA) has been recognized as a promising technique to be used in the fifth generation (5G) networks [1]–[4]. NOMA utilizes the power domain for achieving multiple access, i.e., different users are served at different power levels. Unlike conventional orthogonal MA, such as time-division multiple access (TDMA), NOMA faces strong co-channel interference between different users, and successive interference cancellation (SIC) is used by the NOMA users with better channel conditions for interference management.

The concept of NOMA is essentially a special case of superposition coding developed for broadcast channels (BC). Cover first found the capacity region of a degraded discrete memoryless BC by using superposition coding [5]. Then, the capacity region of the Gaussian BC with single-antenna terminals was established in [6]. Moreover, the capacity region of the multiple-input multiple-output (MIMO) Gaussian BC was found in [7], by applying dirty paper coding (DPC) instead of superposition coding. This paper mainly focuses on the single-antenna scenario.

Specifically, consider a Gaussian BC with a single-antenna transmitter and two single-antenna receivers, where each receiver is corrupted by additive Gaussian noise with unit variance. Denote the ordered channel gains from the transmitter to the two receivers by h_w and h_b, i.e., $|h_w|^2 < |h_b|^2$. For a given channel pair (h_w, h_b), the capacity region is given by [6]

$$
\mathcal{C}^{\mathrm{BC}} \triangleq \bigcup_{a_1+a_2=1,\, a_1,\, a_2 \geq 0} \left\{ (R_1, R_2) : R_1, R_2 \geq 0, \right.
$$
$$
\left. R_1 \leq \log_2\left(1+\frac{a_1 x}{1+a_2 x}\right), R_2 \leq \log_2\left(1+a_2 y\right) \right\}, \quad (1)
$$

where a_i denotes the power allocation coefficient, $x = |h_w|^2\rho$, $y = |h_b|^2\rho$, and ρ denotes the transmit signal-noise-ratio

P. Xu and X. Dai are with Dept. of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei, Anhui, China. Z. Ding and H. V. Poor are with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA. Z.Ding is also with the School of Computing and Communications, Lancaster University, U.K.
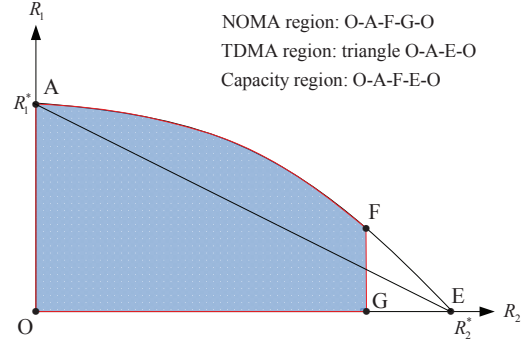

Fig. 1. The capacity region, NOMA and TDMA regions for a given channel pair (h_w, h_b), where the point F is located at $\left(\log_2(1+\frac{y}{2}), \log_2\left(1+\frac{x}{2+x}\right)\right)$.

(SNR). Based on SIC, the rate region achieved by NOMA, denoted by $\mathcal{R}^{\mathrm{N}}$, can be expressed the same as $\mathcal{C}^{\mathrm{BC}}$ in (1), but with an additional constraint $a_1 \geq a_2$ in order to guarantee the quality of service at the user with the poorer channel condition.

In addition, the TDMA region is given by

$$
\mathcal{R}^{\mathrm{T}} \triangleq \left\{ (R_1, R_2) : R_1, R_2 \geq 0, \frac{R_1}{R_1^*} + \frac{R_2}{R_2^*} \leq 1 \right\}, \quad (2)
$$

where $R_1^* = \log_2(1+x)$ and $R_2^* = \log_2(1+y)$.

The three regions are illustrated in Fig. 1. In the rest of this letter, we are interested in the two region boundaries, i.e., the curve A-F and the segment A-E, which represent the optimal rate pairs achieved by NOMA and TDMA, respectively. The relationship between the rate pairs on the two boundaries will be further interpreted based on plane geometry as shown in Section III-A. Then, based on their relationship, the performance of NOMA is characterized in terms of both the sum rate and individual user rates, by considering the conventional TDMA scheme as the benchmark. In a wireless downlink scenario with user pairing, analytical results are developed to demonstrate that NOMA can outperform TDMA when there exists a significant difference between the channel conditions of the scheduled users.

II. PRELIMINARY

Two propositions are provided in this section, which will be used in the next section. Specifically, define $f^N(\cdot)$ and $f^T(\cdot)$ as the following two functions:

$$
f^N(z) = \log_2\left(\frac{(1+x)y}{y+(2^z-1)x}\right), \quad 0 \leq z \leq R_2^*, \quad (3)
$$

$$
f^T(z) = \left(1-\frac{z}{R_2^*}\right) R_1^*, \quad 0 \leq z \leq R_2^*. \quad (4)
$$

For a given $z_0 \in (0, R_2^*)$, two propositions are provided as follows.

Proposition 1: If $z > z_0$, then $f^N(z) + z > f^T(z_0) + z_0$.

Proof: We can first obtain $\frac{d\,f^N}{d\,z} = \frac{-x2^z}{y-x+x2^z}$. Define $f_{sum}^N = f^N + z$, so $\frac{d\,f_{sum}^N}{d\,z} = \frac{y-x}{y-x+x2^z} > 0$, which means that $f_{sum}^N(z)$ is a monotonic increasing function of z. Thus, $f_{sum}^N(z) > f_{sum}^N(z_0) = f^N(z_0) + z_0$.

On the other hand, $f^N(z)$ is a concave function of z (i.e., $\frac{d^2\,f^N}{d\,z^2} < 0$) when $z \in (0, R_2^*)$. Hence

$$\lambda f^N(0) + (1-\lambda)f^N(R_2^*) \le f^N(\lambda \times 0 + (1-\lambda)R_2^*),$$

for $\forall \lambda \in (0,1)$. Since $f^N(0) = R_1^*$ and $f^N(R_2^*) = 0$, we can obtain $f^N(z_0) \ge f^T(z_0)$ by setting $\lambda = 1 - z_0/R_2^*$. This proposition has been proved. ∎

Proposition 2: If $z < z_0$, then $f^N(z) > f^T(z_0)$.

Proof: Since $f^N(z)$ is monotonically decreasing in z, $f^N(z) > f^N(z_0)$ for $z < z_0$. Furthermore, due to the fact that $f^N(z)$ is a concave function of z, $f^N(z_0) > f^T(z_0)$ as discussed above. This proposition has been proved. ∎

III. PERFORMANCE ANALYSIS

In this section, the performance of NOMA will be studied by considering the achievable rates of TDMA as a benchmark.

A. Comparison to TDMA

Here, the individual rates and the sum rate achieved by NOMA will be compared with those of TDMA using plane geometry.

As shown in Fig. 2, for a given channel pair (h_w, h_b) and $|h_w| < |h_b|$, suppose the point N is located at

$$(R_2^N, R_1^N) = \left(\log_2(1 + a_2 y), \log_2\left(1 + \frac{a_1 x}{1 + a_2 x}\right)\right), \quad (5)$$

where $a_1 + a_2 = 1$, $0 \le a_2 \le a_1$; and the point T is located at

$$(R_2^T, R_1^T) = (b_2 R_2^*, b_1 R_1^*), \quad (6)$$

where $b_1 + b_2 = 1$, $b_1, b_2 \ge 0$. This means that the points N and T lie on the curve A-F (NOMA rate pair) and the segment A-E (TDMA rate pair), respectively. In addition, consider three important lines: $R_1 = R_1^N$, $R_2 = R_2^N$ and $R_1 + R_2 = R_1^N + R_2^N$, which represent the two NOMA users' individual rates and their sum rate, respectively. It is easy to prove that $R_1^N + R_2^N < R_2^*$, and these three lines will divide the line segment A-E into four subsegments with intersection points B, C and D.

When considering h_w and h_b to be random variables and fixing a_i, we can define four random events according to the location of point T and these four subsegments as follows.

$$\varepsilon_1 \triangleq \{\text{Point T lies on subsegment A-B}\}, \quad (7)$$
$$\varepsilon_2 \triangleq \{\text{Point T lies on subsegment B-C}\}, \quad (8)$$
$$\varepsilon_3 \triangleq \{\text{Point T lies on subsegment C-D}\}, \quad (9)$$
$$\varepsilon_4 \triangleq \{\text{Point T lies on subsegment D-E}\}. \quad (10)$$

These events comprehensively reflect the relationship between the rates (including the individual rates and the sum rate) of NOMA and TDMA, i.e.,

$$\varepsilon_1 = \{R_1^N < R_1^T, R_2^N > R_2^T, R_1^N + R_2^N > R_1^T + R_2^T\}, \quad (11)$$
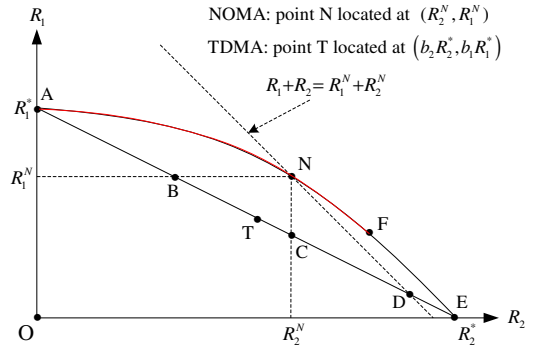$$\varepsilon_2 = \{R_1^N > R_1^T, R_2^N > R_2^T, R_1^N + R_2^N > R_1^T + R_2^T\}, \quad (12)$$

Fig. 2. Comparison of the rate pairs achieved by NOMA and TDMA schemes for a given channel pair (h_w, h_b), where (R_2^N, R_1^N) is defined in (5). In this example, T lies on the line segment B-C.

$$\varepsilon_3 = \{R_1^N > R_1^T, R_2^N < R_2^T, R_1^N + R_2^N > R_1^T + R_2^T\}, \quad (13)$$
$$\varepsilon_4 = \{R_1^N > R_1^T, R_2^N < R_2^T, R_1^N + R_2^N < R_1^T + R_2^T\}. \quad (14)$$

Notice that $(R_1^N, R_2^N, R_1^T, R_2^T)$ satisfies the relationship $R_1^N = f^N(R_2^N)$ and $R_1^T = f^T(R_2^T)$ as shown in (3) and (4). Hence, based on Propositions 1 and 2 by replacing (f^N, z, f^T, z_0) with $(R_1^N, R_2^N, R_1^T, R_2^T)$, we can remove some redundant conditions for each event, i.e.,

$$\varepsilon_1 \overset{(a)}{=} \{R_1^N < R_1^T, R_2^N > R_2^T\}, \quad (15)$$
$$\varepsilon_2 \overset{(b)}{=} \{R_1^N > R_1^T, R_2^N > R_2^T\}, \quad (16)$$
$$\varepsilon_3 \overset{(c)}{=} \{R_2^N < R_2^T, R_1^N + R_2^N > R_1^T + R_2^T\}, \quad (17)$$
$$\varepsilon_4 \overset{(d)}{=} \{R_1^N + R_2^N < R_1^T + R_2^T\}, \quad (18)$$

where (a) and (b) are based on Proposition 1; (c) is based on Proposition 2; and (d) is based on the converse-negative proposition of Proposition 1 (i.e., $\{R_1^N + R_2^N < R_1^T + R_2^T\} \Rightarrow \{R_2^N < R_2^T\}$) and Proposition 2.

Remark 1: Among these four events, of particular interest is ε_2 which represents the situation in which NOMA outperforms TDMA in terms of not only the sum rate but also each individual rate.

In the next subsection, the probability of each event will be calculated, which characterizes the performance of NOMA in comparison with TDMA.

B. Probability Analysis

Let the two users in the considered BC be selected from M mobile users in a downlink communication scenario, as motivated in [8]. Without loss of generality, assume that all the users' channels are ordered as $|h_1|^2 \le \cdots \le |h_M|^2$, where h_m is the Rayleigh fading channel gain from the base station to the m-th user. Considered that the m-th user is paired with the n-th user to perform NOMA. Hence $h_w = h_m$, $h_b = h_n$, and x and y can be rewritten as $x = \rho|h_m|^2$, $y = \rho|h_m|^2$, with joint probability density function (PDF) as follows [9]:

$$f_{X,Y}(x,y) = w_1 f(x)f(y)[F(x)]^{m-1}[1-F(y)]^{M-n}$$
$$\times [F(y) - F(x)]^{n-1-m}, \quad 0 < x < y, \quad (19)$$

where $f(x) = \frac{1}{\rho}e^{-\frac{x}{\rho}}$, $F(x) = 1 - e^{-\frac{x}{\rho}}$, and $w_1 = \frac{M!}{(m-1)!(n-1-m)!(M-n)!}$. A fixed power allocation strategy

(a_1, a_2) is considered in this NOMA system for the sake of simplicity. Dynamically changing (a_1, a_2) according to the random channel state information (CSI) could achieve a larger ergodic rate region [10], but at the expense of higher complexity.

Using the PDF of x and y, the probability of each event defined in the previous subsection can be calculated in order to evaluate the performance of NOMA. The probability of the event ε_2 is first given in the following lemma, where we set $b_2 = 1/2$ (each user is allocated an equal-length time slot, which is also called "naive TDMA" in [11]) for simplicity.

Lemma 1: Given (M, m, n, ρ, a_2) and $b_2 = 1/2$, the probability that NOMA achieves larger individual rates than conventional TDMA for both user m and user n is given by

$$P(\varepsilon_2) = w_1 \sum_{k=0}^{m-1} (-1)^{m-1-k} C_{m-1}^k \left[\sum_{i=0}^{n-1} \frac{(-1)^{n-1-i} C_{n-1}^i d^{M-i}}{M-i} - \sum_{i=0}^{k} \sum_{j=0}^{n-1-k} \frac{(-1)^{n-1-i-j} C_k^i C_{n-1-k}^j d^{M-i}}{M-i-j} \right]. \quad (20)$$

Proof: From (16), $P(\varepsilon_2)$ can be calculated as follows:

$$P(\varepsilon_2) = P(R_1^N > R_1^*/2, R_2^N > R_2^*/2)$$
$$\overset{(a)}{=} P(x < w_2, y > w_2)$$
$$= w_1 \int_{w_2}^{+\infty} f(y)[1-F(y)]^{M-n}$$
$$\times \left(\int_0^{w_2} f(x)[F(x)]^{m-1}[F(y)-F(x)]^{n-1-m} dx \right) dy$$
$$\overset{(b)}{=} w_1 \int_{w_2}^{+\infty} f(y)[1-F(y)]^{M-n}$$
$$\times \left(\int_d^1 (1-t)^{m-1}(t - e^{-\frac{y}{\rho}})^{n-1-m} dt \right) dy$$
$$= w_1 \int_{w_2}^{+\infty} f(y)[1-F(y)]^{M-n}$$
$$\times \left(\int_{d-e^{-\frac{y}{\rho}}}^{1-e^{-\frac{y}{\rho}}} (1 - e^{-\frac{y}{\rho}} - u)^{m-1} u^{n-1-m} du \right) dy$$
$$\overset{(c)}{=} w_1 \int_{w_2}^{+\infty} f(y)[1-F(y)]^{M-n} \sum_{k=0}^{m-1} (-1)^{m-1-k} C_{m-1}^k$$
$$\times \frac{(1-e^{-\frac{y}{\rho}})^{n-1} - (1-e^{-\frac{y}{\rho}})^k(d-e^{-\frac{y}{\rho}})^{n-1-k}}{n-1-k} dy$$
$$= w_1 \sum_{k=0}^{m-1} \frac{(-1)^{m-1-k} C_{m-1}^k}{n-1-k} \left[\underbrace{\int_0^d v^{M-n}(1-v)^{n-1} dv}_{Q_1} - \underbrace{\int_0^d v^{M-n}(1-v)^k(d-v)^{n-1-k} dv}_{Q_{2,k}} \right] \quad (21)$$

where (a) follows the definition $w_2 = \frac{1-2a_2}{a_2^2}$; (b) follows $d = e^{-\frac{w_2}{\rho}}$; and (c) follows $C_p^q = \frac{p!}{q!(p-q)!}, p > q$. Furthermore,

the two terms Q_1 and $Q_{2,k}$ can be calculated as follows:

$$Q_1 = \sum_{i=0}^{n-1} (-1)^{n-1-i} C_{n-1}^i \int_0^d v^{M-1-i} dv$$
$$= \sum_{i=0}^{n-1} \frac{(-1)^{n-1-i} C_{n-1}^i d^{M-i}}{M-i}, \quad (22)$$

$$Q_{2,k} = \sum_{i=0}^{k} \sum_{j=0}^{n-1-k} (-1)^{n-1-i-j} C_k^i C_{n-1-k}^j d^j \int_0^d v^{M-1-i-j} dv$$
$$= \sum_{i=0}^{k} \sum_{j=0}^{n-1-k} \frac{(-1)^{n-1-i-j} C_k^i C_{n-1-k}^j d^{M-i}}{M-i-j}. \quad (23)$$

Substituting the above two relationships into (21), this lemma has been proved. ∎

Moreover, for the first event, it is not difficult to obtain that

$$P(\varepsilon_1) = P(R_2^N > R_2^T) - P(\varepsilon_2)$$
$$= 1 - w_3 \sum_{i=0}^{n-1} \frac{(-1)^i C_{n-1}^i}{M-n+i+1} \left(1 - d^{M-n+i+1} \right) - P(\varepsilon_2), \quad (24)$$

where $w_3 = \frac{M!}{(m-1)!(M-m)!}$. For the fourth event, from [8] (Theorem 1), we have

$$P(\varepsilon_4) = P(R_1^N + R_2^N < R_1^T + R_2^T) =$$
$$1 - w_1 \sum_{i=0}^{n-1-m} \frac{(-1)^i C_{n-1-m}^i}{m+i} \int_{\sqrt{w_2+1}-1}^{w_2} f(y)[F(y)]^{n-1-m-i}$$
$$\times [1-F(y)]^{M-n} \left([F(y)]^{m+i} - \left[F\left(\frac{w_2 - y}{1+y} \right) \right]^{m+i} \right) dy$$
$$- w_3 \sum_{j=0}^{n-1} \frac{(-1)^j C_{n-1}^j}{M-n+j+1} d^{-(M-n+j+1)}. \quad (25)$$

Thus, the probability of the third event can be written as

$$P(\varepsilon_3) = 1 - P(\varepsilon_1) - P(\varepsilon_2) - P(\varepsilon_4). \quad (26)$$

Now, $P(\varepsilon_i)$, $i = 2, 1, 4, 3$, have been obtained as in Eqs. (21), (24), (25), and (26), respectively.

Special Case: The expression for $P(\varepsilon_2)$ in Lemma 1 can be simplified when considering a special pairing case, i.e., $m = 1$, $n = M$. In this case, $k = 0$, and Q_1 and $Q_{2,0}$ in (22) and (23), respectively, can be derived as

$$Q_1 = -\frac{1}{M} \sum_{i=0}^{M-1} (-1)^{M-i} C_M^i d^{M-i}$$
$$= -\frac{1}{M} \left[\sum_{i=0}^{M} C_M^i (-d)^{M-i} - 1 \right] = \frac{1 - (1-d)^M}{M}, \quad (27)$$

$$Q_{2,0} = \sum_{j=0}^{M-1} \frac{(-1)^{M-1-j} C_{M-1}^j d^M}{M-j} = -\frac{d^M}{M} \sum_{i=0}^{M-1} (-1)^{M-j} C_M^j$$
$$= -\frac{d^M}{M} \left[\sum_{i=0}^{M} (-1)^{M-j} C_M^j - 1 \right]$$
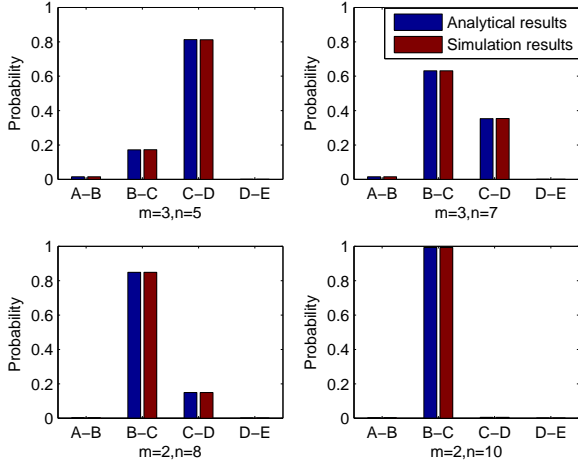$$= -\frac{d^M}{M} \left[(1-1)^M - 1 \right] = \frac{d^M}{M}. \quad (28)$$

Fig. 3. Probability that the point T lies on a certain segment in Fig. 2, where $\rho = 25$ (dB), $a_2 = 1/\sqrt{\rho}$.

Hence, $P(\varepsilon_2) = 1 - (1-d)^M - d^M$. The optimal d in this case is $1/2$, which implies $a_2 = \frac{-1+\sqrt{1+\rho\ln 2}}{\rho \ln 2} < \frac{1}{2}$, and $P(\varepsilon_2) = 1 - \frac{1}{2^{M-1}}$. Here $\ln(\cdot)$ denotes the natural logarithm.

Remark 2: This special case shows that $P(\varepsilon_2) \to 1$ when M is sufficiently large. This means that, almost for all the possible channel realizations, NOMA achieves larger individual rates than naive TDMA for both user m and n as long as the difference between the better and worse channel gains is sufficiently large. This phenomenon is also valid for other pairing cases (i.e., $(n,m) \neq (M,1)$) as verified via some numerical examples in the next section.

IV. NUMERICAL RESULTS

In this section, the performance of NOMA is evaluated in comparison with TDMA by using computer simulations. The total number of users in the wireless downlink system is $M = 10$, and different choices of (m,n) will be considered.

In Fig. 3, the probability of each event defined in Section III-A is displayed via column diagrams. Specifically, the probabilities that the point T lies on subsegments A-B, B-C, C-D and D-E in Fig. 2 are displayed, where we set $\rho = 25$ dB and $a_2 = 1/\sqrt{\rho}$ for simplicity. Four different user pairs (m,n) are considered, which shows that the probability that the point T lies on subsegment B-C (i.e., $P(\varepsilon_2)$) increases with the value of $(n - m)$, as discussed in Remark 2. In Fig. 4, additional numerical results are provided to show $P(\varepsilon_2)$ as a function of n. As shown in this figure, $P(\varepsilon_2)$ increases with the value of $(n - m)$, i.e., NOMA is prone to perform better than TDMA in terms of individual rates when the difference between the users' channels is large. In addition, it is worth pointing out that the Monte Carlo simulation results provided in Figs. 3 and 4 match well with the analytical results developed in (21), (24), (25) and (26). In Fig. 5, individual rates of NOMA and TDMA averaged over the fading channels are depicted as functions of SNR (i.e., ρ), where we set $(m,n) = (1,M)$, and $a_2 = (\sqrt{\rho\ln(2) + 1} - 1)/(\rho\ln(2))$ according to the special case in Section III-B. As shown in this figure, NOMA has
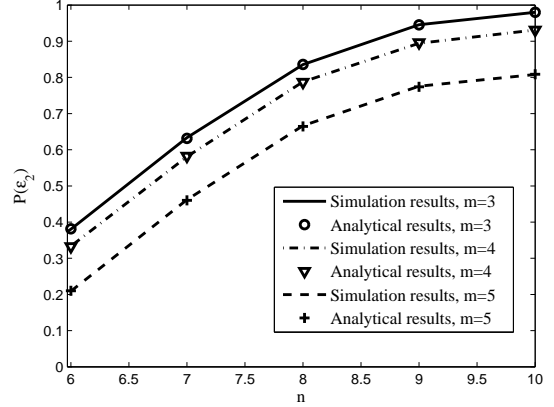

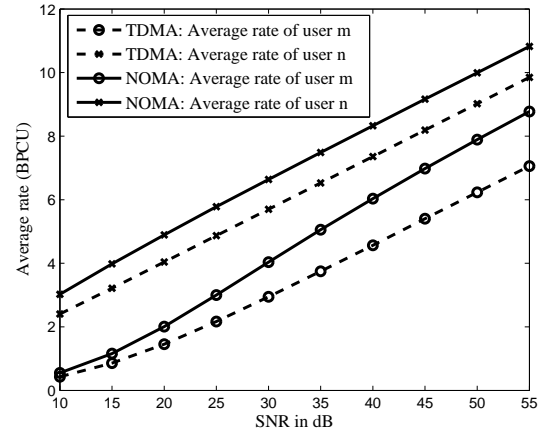Fig. 4. Probability of the event ε_2, where $\rho = 25$ (dB), $a_2 = 1/\sqrt{\rho}$.

Fig. 5. Average individual rates, $a_2 = \frac{\sqrt{\rho\ln(2)+1}-1}{\rho\ln(2)}$, $m = 1$, $n = 10$.

a constant performance gain over TDMA for each individual rate. When $\rho = 55$ dB, the performance gains with respect to user m and user n are about 1 bits per channel uses (BPCU) and 2 BPCU, respectively. This is due to the fact that $P(\varepsilon_2) \to 1$ in this case, i.e., NOMA outperforms TDMA in terms of each user's rate for almost all the possible realizations of (h_m, h_n).

V. CONCLUSIONS

This letter has investigated the performance of NOMA in a downlink network from an information theoretic perspective. The relationship among the BC capacity region, the NOMA rate region and the TDMA rate region was first described. According to their relationship, the performance of NOMA was evaluated in terms of both the sum rate and users' individual rate, by considering TDMA as the benchmark. Future work of interest is to dynamically change power allocation according to instantaneous CSI for enlarging the ergodic achievable rates [10]. Moreover, it is important to establish the connection between MIMO NOMA and information theoretic MIMO broadcasting concepts.

VI. ACKNOWLEDGMENT

REFERENCES

[1] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *Proc. IEEE Intl. Symp. Personal, Indoor and Mobile Radio Communications (PIMRC)*, London, U.K., Sep. 2013, pp. 611–615.

[2] Q. Li, H. Niu, A. Papathanassiou, and G. Wu, "5G network capacity: Key elements and technologies," *IEEE Veh. Technol. Mag.*, vol. 9, no. 1, pp. 71–78, Mar. 2014.

[3] 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 Processing Lett.*, vol. 21, no. 12, pp. 1501–1505, Sep. 2014.

[4] S. Timotheou and I. Krikidis, "Fairness for non-orthogonal multiple access in 5G systems," *IEEE Signal Processing Lett. (Accepted)*, 2015.

[5] T. Cover, "Broadcast channels," *IEEE Tras. Inf. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.

[6] P. Bergmans, "A simple converse for broadcast channels with additive white gaussian noise (corresp.)," *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 279–280, Mar. 1974.

[7] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[8] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G non-orthogonal multiple access," *arXiv preprint arXiv:1412.2799*, 2014.

[9] H. A. David and H. N. Nagaraja, *Order statistics*, 3rd ed. New York, NY, USA: Wiley, 2003.

[10] L. Li and A. J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels - Part I: Ergodic capacity," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 1083–1102, Mar. 2001.

[11] T. Cover and J. Thomas, *Elements of information theory, second edition*. Wiley-Interscience: NJ, 2006.