Performance Analysis of NOMA Uplink Networks under Statistical QoS Delay Constraints

Mouktar Bello* , Wenjuan Yu †, Arsenia Chorti* and Leila Musavian§

*ETIS UMR8051, CY Universit, ENSEA, CNRS, F-95000, Cergy, France
†5GIC, Institute of Communication Systems, University of Surrey, Guildford, GU2 7XH, UK
§School of Computer Science and Electronic Engineering, University of Essex, Colchester, CO4 3SQ, UK

Abstract—**In the fifth generation and beyond (B5G), delay constraints emerge as a topic of particular interest, e.g. for ultra-reliable low latency communications (URLLC) such as autonomous vehicles and enhanced reality. In this paper, we study the performance of a two-user uplink NOMA network under statistical quality of service (QoS) delay constraints, captured through each user's effective capacity (EC). We propose novel closed-form expressions for the EC of the NOMA users and show that in the high signal to noise ratio (SNR) region, the "strong" NOMA user has a limited EC, assuming the same delay constraint as the "weak" user. We demonstrate that for the weak user, OMA achieves higher EC than NOMA at small values of the transmit SNR, while NOMA outperforms OMA in terms of EC at high SNRs. On the other hand, for the strong user the opposite is true, i.e., NOMA achieves higher EC than OMA at small SNRs, while OMA becomes more beneficial at high SNRs. This result raises the question of introducing "adaptive" OMA / NOMA policies, based jointly on the users' delay constraints as well as on the available transmit power.**

Index Terms—**NOMA, QoS, low latency, effective capacity, B5G.**

I. INTRODUCTION

Non-orthogonal multiple access (NOMA) schemes have attracted a lot of attention recently, allowing multiple users to be served simultaneously with enhanced spectral efficiency; it is known that the boundary of achievable rate pairs (in the case of two users) using NOMA is outside the capacity region achievable with orthogonal multiple access (OMA) techniques [1] or other schemes [2]. Superior achievable rates are attainable through the use of superposition coding at the transmitter and of successive interference cancellation (SIC) at the receiver [3], [4]. The SIC receiver decodes multi-user signals with descending received signal power and subtracts the decoded signal(s) from the received superimposed signal, so as to improve the signal-to-interference ratio. The process is repeated until the signal of interest is decoded. In uplink NOMA networks, the strongest user's signal is decoded first (as opposed to downlink NOMA networks in which the inverse order is applied).

Besides, in a number of emerging applications, delay quality of service (QoS) becomes increasingly important, e.g., ultra reliable low latency communication (URLLC) systems. Furthermore, in future wireless networks, users are expected to necessitate flexible delay guarantees for achieving different service requirements. In order to satisfy diverse delay require-ments, a simple and flexible delay QoS model is imperative to be applied and investigated. In this respect, the effective capacity (EC) theory can be employed [5], [6] [7], with EC denoting the maximum constant arrival rate which can be served by a given service process, while guaranteeing the required statistical delay provisioning. We studied the delay-constrained communications for a downlink NOMA network in [4] and with secrecy constraints [8] in [9]. The present analysis complements [4], focusing on uplink transmissions. NOMA, as a more spectrum-efficient technique, is considered to be promising for supporting the massive number of devices to access the uplink connections. Hence, we believe that it is important to investigate the delay-constrained achievable rate for an uplink NOMA network.

In this paper, we provide a performance evaluation of the uplink transmission for a two-user NOMA network under delay constraints, captured through the users' ECs. We note that the EC is a QoS aware data-link layer metric [6], that captures the achievable rate under a delay violation probability threshold.

In this work, we first derive novel closed-form expressions for the ECs of both users; we then provide four Lemmas for the asymptotic performance of the network with NOMA and OMA. The conclusions drawn are supported by an extensive set of simulations. The paper is organized as follows. In Section II we investigate the EC of a two user uplink NOMA system under the delay QoS constraints. Simulation results are given in Section III, followed by conclusions in Section IV.

II. EFFECTIVE CAPACITY OF TWO-USER NOMA UPLINK NETWORK

Assume a two-user NOMA uplink network with users U_1 and U_2 in a Rayleigh block fading propagation channel, with respective channel gains during a transmission block denoted by $|h_1|^2 < |h_2|^2$. The users transmit corresponding symbols S_1, S_2 respectively, with power $\mathbb{E}[|S_i|^2] = P_i, i = 1, 2$ and the total power $P_T = \sum_{i=1}^{2} P_i = 1$. Here, P_i is the power coefficient for the user i and normalized transmission powers are assumed [10]. The received superimposed signal can be expressed as [11]

$$Z = \sum_{i=1}^{2} \sqrt{P_i} h_i S_i + w, \qquad (1)$$

where w denotes a zero mean circularly symmetric complex Gaussian random variable with variance σ^2. The receiver will first decode the symbol of the strong user treating the transmission of the weak as interference. After decoding it, the receiver will suppress it from Z and decode the signal of the weak user. Following the SIC principle and denoting by $\rho = \frac{1}{\sigma^2}$ the transmit SNR, the achievable rates, in b/s/Hz, for user $U_i, i = 1, 2$, is expressed as: [12]

$$R_i = \log_2 \left[1 + \frac{\rho P_i |h_i|^2}{1 + \rho \sum_{l=1}^{i-1} P_l |h_l|^2} \right]. \quad (2)$$

To clarify further, let θ_i be the statistical delay QoS exponent of the i-th user, and assume that the service process satisfies the Gärtner-Ellis theorem [6]. The delay exponent θ_i captures how strict the delay constraint is [6]. A slower decay rate can be represented by a smaller θ_i, which indicates that the system is more delay tolerant, while a larger θ_i corresponds to a system with more stringent QoS requirements. Applying the EC theory in a uplink NOMA with two users, the i-th user's EC over a block-fading channel, is defined as:

$$E_c^i = -\frac{1}{\theta_i T_f B} \ln \left(\mathbb{E}\left[e^{-\theta_i T_f B R_i} \right] \right) \quad \text{(in b/s/Hz)}, \quad (3)$$

where T_f is the fading-block length, B is the bandwidth and $\mathbb{E}[\cdot]$ denotes expectation over the channel gains. By inserting R_i into (3), we obtain the following expression for the EC of the i-th user

$$E_c^i = \frac{1}{\beta_i} \log_2 \left(\mathbb{E}\left[(1 + \frac{\rho P_i |h_i|^2}{1 + \rho \sum_{l=1}^{i-1} P_l |h_l|^2})^{\beta_i} \right] \right) \quad (4)$$

where $\beta_i = -\frac{\theta_i T_f B}{\ln 2}, i = 1, 2$, is the normalized (negative) QoS exponent.

A. ECs in a Two-user NOMA Uplink Network

For the ordering of the channel gains we make use of the theory of order statistics in the following analysis [13].

Assuming a Rayleigh wireless environment, the channel gains, denoted by $x_i = |h_i|^2, i = 1, 2$, are exponentially distributed with probability density function (PDF) and cumulative density function (CDF) respectively given by $f(x_i) = e^{-x_i}$, $F(x_i) = 1 - e^{-x_i}$.

Then, according to order statistics [13], the ordered channel gains have respective PDFs $f_{i:2}(x_i), i = 1, 2$, and joint PDF $f(x_1, x_2)$ that are expressed as

$$f_{1:2}(x_1) = 2e^{-2x_1}, \quad (5)$$
$$f_{2:2}(x_2) = 2e^{-x_2} \left(1 - e^{-x_2} \right), \quad (6)$$
$$f(x_1, x_2) = 2e^{-x_1} e^{-x_2}. \quad (7)$$

As a result, the EC of User 1, denoted by E_c^1 is expressed as

$$\begin{aligned}
E_c^1 &= \frac{1}{\beta_1} \log_2 (\mathbb{E}[(1 + \rho P_1 x_1)^{\beta_1}]) \\
&= \frac{1}{\beta_1} \log_2 \left(\int_0^\infty (1 + \rho P_1 x_1)^{\beta_1} f_{1:2}(x_1) dx_1 \right) \\
&= \frac{1}{\beta_1} \log_2 \left(\frac{2}{P_1 \rho} \times U \left(1, 2 + \beta_1, \frac{2}{\rho P_1} \right) \right). \quad (8)
\end{aligned}$$

where $U(\cdot, \cdot, \cdot)$ denotes the confluent hypergeometric function [4]. On the other hand, the EC of the User 2 is evaluated as

$$\begin{aligned}
E_c^2 &= \frac{1}{\beta_2} \log_2 \left(\mathbb{E}\left[\left(1 + \frac{\rho P_2 x_2}{1 + \rho P_1 x_1} \right)^{\beta_2} \right] \right) \\
&= \frac{1}{\beta_2} \log_2 \left(\int_0^\infty \int_{x_1}^\infty \left(1 + \frac{\rho P_2 x_2}{1 + \rho P_1 x_1} \right)^{\beta_2} f(x_1, x_2) dx_2 dx_1 \right) \\
&= \frac{1}{\beta_2} \log_2 \left(2P_2^{1-\beta_2} (\rho P_2)^{\beta_2} e^{\frac{1}{\rho P_2}} e^{-\frac{(P_1 - P_2)}{\rho P_2}} \right) \\
&\quad + \frac{1}{\beta_2} \log_2 \left(\sum_{j=0}^{-\beta_2} \binom{-\beta_2}{j} (\rho P_1)^j \times \sum_{k=0}^\infty \frac{(-1)^k (P_2 - P_1)^k}{k!\,(1 + j + k)} \right. \\
&\quad \times \left[\Gamma[2 + \beta_2 + j + k, \frac{1}{\rho P_2}] \right. \\
&\quad \left. \left. - (\rho P_2)^{-1-j-k} \Gamma[1 + \beta_2, \frac{1}{\rho P_2}] \right] \right), \quad (9)
\end{aligned}$$

with $\Gamma(\cdot, \cdot)$ denoting the incomplete Gamma function [4].

The proof for deriving E_c^1 is omitted due to space limitations while for E_c^2 is provided in Appendix I.

In order to perform a comparative performance analysis, here we provide the achievable data rates for a two-user OMA network, denoted by $\widetilde{R}_i, i = 1, 2$, given as

$$\widetilde{R}_i = \frac{1}{2} \log_2 \left(1 + \rho P_T |h_i|^2 \right), i = 1, 2 \quad (10)$$

Note that $\frac{1}{2}$ is due to the equal allocation of resources to both users. The corresponding expressions are obtained for the ECs of both users in a OMA network, denoted by $\widetilde{E}_c^i$, given as

$$\widetilde{E}_c^i = \frac{1}{\beta_i} \log_2 \left(\mathbb{E}\left[(1 + \rho P_T |h_i|^2)^{\frac{\beta_i}{2}} \right] \right)$$

$i.e,$ (11)

$$\widetilde{E}_c^1 = \frac{1}{\beta_1} \log_2 \left(\frac{2}{\rho} \times U \left(1, 2 + \frac{\beta_1}{2}, \frac{2}{\rho} \right) \right)$$

$$\widetilde{E}_c^2 = \frac{1}{\beta_2} \log_2 \left(\frac{2}{\rho} \sum_{k=0}^1 \binom{1}{k} (-1)^k \times U \left(1, 2 + \frac{\beta_2}{2}, \frac{1+k}{\rho} \right) \right)$$

The proof is omitted due to space limitations.

B. Asymptotic Analysis

We first perform an asymptotic analysis with respect to the SNR. Our results are summarized in Lemma 1.

Lemma 1: *In the low and high SNR regimes, respectively, the following conclusions hold:*

1) *When $\rho \to 0$, then, $E_c^1 \to 0$, $E_c^2 \to 0$, $\widetilde{E}_c^1 \to 0$, $\widetilde{E}_c^2 \to 0$, $E_c^1 - \widetilde{E}_c^1 \to 0$, $E_c^2 - \widetilde{E}_c^2 \to 0$;*

2) *When $\rho \to +\infty$, then $E_c^1 \to +\infty$, $E_c^2 \to \frac{1}{\beta_2} \log_2 \left(\mathbb{E}\left[\left(1 + \frac{P_2 |h_2|^2}{P_1 |h_1|^2} \right)^{\beta_2} \right] \right)$, $\widetilde{E}_c^1 \to +\infty$, $\widetilde{E}_c^2 \to +\infty$, $E_c^1 - \widetilde{E}_c^1 \to +\infty$, $E_c^2 - \widetilde{E}_c^2 \to -\infty$.*

Proof: : The proof is provided in Appendix II. ∎

Lemma 1 indicates that the ECs of both users are vanishingly small at low values of ρ, irrespective of employing NOMA or OMA. On the other hand, at high SNRs, we notice

that the EC of the strong user with NOMA is limited to a finite value. On the contrary, for the weaker user, when $\rho >> 1$, its achievable EC in the NOMA uplink increases without bound. This is the exact opposite of the downlink scenario, where it is the weaker user which is limited in terms of EC, when $\rho >> 1$ [4].

Now, the question is how the ECs evolve with ρ between the two asymptotic regimes. To answer this question and to further analyze the impact of ρ on the individual EC, we look at the derivatives with the respect of ρ [4].

Lemma 2: *For the EC of the User 1, in a two-user uplink network the following hold:*

1) $\frac{\partial E_c^1}{\partial \rho} \geq 0$ *and* $\frac{\partial \widetilde{E}_c^1}{\partial \rho} \geq 0$, $\forall \rho$;

2) *When* $\rho \to 0$, *then* $\lim_{\rho \to 0} \left(\frac{\partial (E_c^1 - \widetilde{E}_c^1)}{\partial \rho} \right) = \frac{P_1 - \frac{1}{2}}{\ln 2} \mathbb{E}[|h_1|^2]$;

3) *When* $\rho >> 1$, *then* $\frac{\partial (E_c^1 - \widetilde{E}_c^1)}{\partial \rho} \approx \frac{1}{2\rho \ln 2} \geq 0$ *and it approaches 0 when* $\rho \to \infty$.

    Proof: : The proof is provided in Appendix III. ∎

Lemma 2 indicates that for User 1, when the transmit SNR ρ is very small, the EC with OMA increases faster than the EC with NOMA. On the other hand, Lemma 2 shows that when the transmit SNR is very large, the EC with NOMA increases faster than with OMA.

Combining Lemma 2 and Lemma 1, we can conclude that, $E_c^1 - \widetilde{E}_c^1$ starts at vanishingly small value, first decreases, and subsequently increases to ∞ at a gradually reducing speed. This means that for the weaker user, OMA achieves higher EC than NOMA at small values of the transmit SNR ρ. At high values of ρ, NOMA becomes more beneficial for the weak user. Finally, when $\rho \to \infty$ the performance gain of NOMA over OMA reaches a constant value in the case of User 1.

Lemma 3: *For the EC of the User 2, in a two-user uplink network the following hold:*

1) $\frac{\partial E_c^2}{\partial \rho} \geq 0$ *and* $\frac{\partial \widetilde{E}_c^2}{\partial \rho} \geq 0$, $\forall \rho$;

2) *When* $\rho \to 0$, *then* $\lim_{\rho \to 0} \left(\frac{\partial (E_c^2 - \widetilde{E}_c^2)}{\partial \rho} \right) = \frac{P_2}{2 \ln 2} \mathbb{E}[|h_2|^2]$

3) *When* $\rho >> 1$, *then* $\frac{\partial (E_c^2 - \widetilde{E}_c^2)}{\partial \rho} \approx -\frac{1}{2 \ln 2} \frac{1}{\rho} < 0$ *and it approaches 0 when* $\rho \to \infty$.

    Proof: The proof is provided in Appendix IV. ∎

Lemma 3 indicates that, for User 2, when the transmit SNR ρ is very small, the uplink EC with NOMA increases faster than that with OMA. On the other hand, when the transmit SNR is very large, the uplink EC with OMA increases faster than that with NOMA. Combining Lemma 3 and Lemma 1, we can conclude that, $E_c^2 - \widetilde{E}_c^2$ starts at an initial vanishingly small value, first increases, and subsequently decreases to $-\infty$ with a gradually diminishing rate. This means that for the stronger user, NOMA achieves higher EC than OMA at small values of the transmit SNR ρ. At high values of ρ, OMA becomes more beneficial for the strong user. Finally, when $\rho \to \infty$ the performance gain of OMA over NOMA reaches a constant value, for the stronger user.

Finally, we investigate the sum ECs when using OMA and NOMA, denoted by V_N and V_O,

$$V_N = E_c^1 + E_c^2, \quad (12)$$
$$V_O = \widetilde{E}_c^1 + \widetilde{E}_c^2. \quad (13)$$

Our conclusions are drawn in Lemma 4.

Lemma 4: *For the sum EC with NOMA, denoted by V_N, and with OMA, denoted by V_O, in a two-user uplink network, the following hold:*

1) $\frac{\partial V_N}{\partial \rho} \geq 0$ *and* $\frac{\partial V_O}{\partial \rho} \geq 0$, $\forall \rho$;

2) *When* $\rho \to 0$, $V_N \to 0$, $\lim_{\rho \to 0} \left(\frac{\partial V_N}{\partial \rho} \right) = \frac{P_1}{\ln 2} \mathbb{E}[|h_1|^2] + \frac{P_2}{\ln 2} \mathbb{E}[|h_2|^2] \geq 0$, *and* $V_O \to 0$, $\lim_{\rho \to 0} \left(\frac{\partial V_O}{\partial \rho} \right) = \frac{P_1}{2 \ln 2} \mathbb{E}[|h_1|^2] + \frac{P_2}{2 \ln 2} \mathbb{E}[|h_2|^2] \geq 0$;

3) *When* $\rho >> 1$, $V_N \to \infty$, $\lim_{\rho \to \infty} \left(\frac{\partial V_N}{\partial \rho} \right) = 0$, *and* $V_O \to \infty$, $\lim_{\rho \to \infty} \left(\frac{\partial V_O}{\partial \rho} \right) = 0$.

    Proof: The proof is provided in Appendix V. ∎

Lemma 4 indicates that when NOMA is applied, the sum EC has a constant increasing rate at small value of the transmit SNR ρ that depends on the average of the channel power gains and the allocated power coefficients. A similar conclusion is reached when using OMA. On the other hand, when $\rho >> 1$, Lemma 4 indicates that the rate at which the sum ECs increase reaches a plateau, both in the case of NOMA and OMA.

III. NUMERICAL RESULTS

In this section, the Lemmas presented in Section II will be validated through Monte Carlo simulations. We consider a two user uplink NOMA system, with the following settings: normalized transmission powers for both users, $P_1 = 0.2$, $P_2 = 0.8$, normalized delay exponent $\beta_1 = \beta_2 = -1$ for both users, unless otherwise stated.

In Fig. 1 the ECs of the two-user uplink NOMA and OMA networks are depicted versus the transmit SNR. We note that for the weak user, OMA is advantageous than NOMA for low transmit SNRs, and NOMA is advantageous OMA at high transmit SNRs. Reverse conclusions can be drawn for the strong user. We notice also that the EC of the strong user converges at high SNRs. This provides numerical validation for Lemma 1.

Figs. 2 and 3, show respectively the EC of User 1 and User 2, versus the transmit SNR, for different values of $\beta_1 = \beta_2 = \beta$. When the delay constraints become more stringent, i.e., β decreases (equivalently, θ increases), the individual link-layer rates in NOMA decrease, for both users.

In Fig. 4, the ECs of the strong and weak users are depicted across different SNR values, ($\rho \in \{1, 10, 30, 40, 50\}$ dB, as functions of the (negative) normalized delay exponent, for NOMA and OMA scenarios. We notice that the EC of each user is identical for NOMA and OMA, for small and large values of the normalized delay exponent. And with the transmit SNR ρ increasing, the EC increases for both users.

Fig. 5 shows the difference of the EC in NOMA and the EC in OMA of the weak user. This curve starts initially at zero,
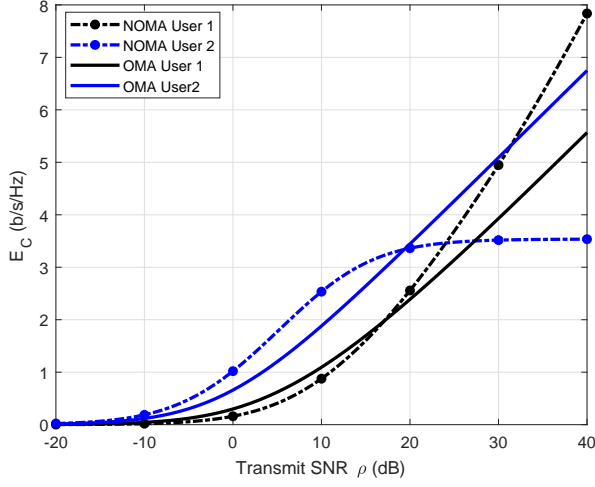
Fig. 1. E_c^1, E_c^2 in a two-user NOMA uplink network compared to Ecs of two users OMA, versus ρ
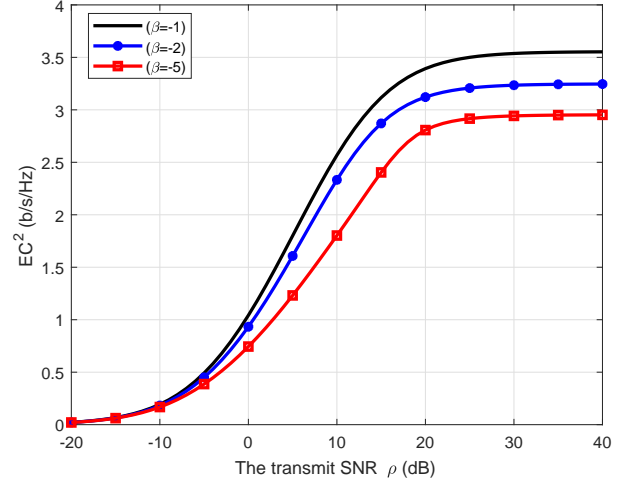

Fig. 3. E_c^2 versus the transmit SNR ρ for several delays.
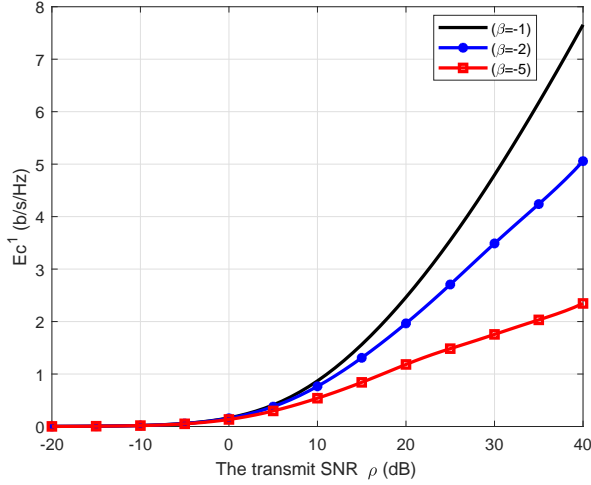

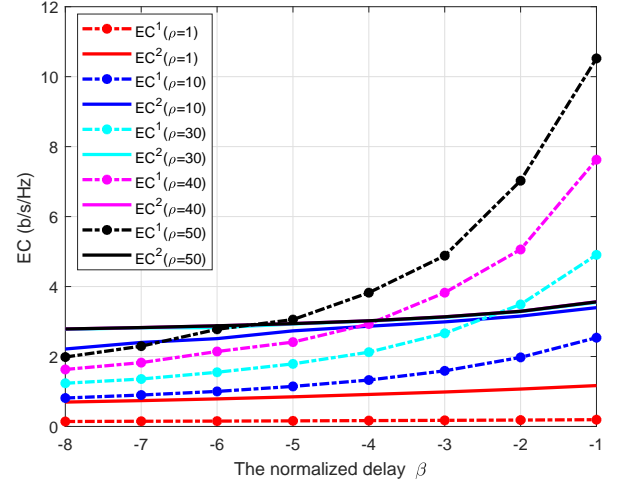Fig. 2. E_c^1 versus the transmit SNR, for several delays.

Fig. 4. E_c^1 and E_c^2 in a two-user NOMA compared to ECs of two users OMA, versus normalized delay β, for different values of ρ.

then decreases to a certain minimum and starts increasing at the high values of transmit SNR. This confirms Lemma 2. When the delay is equal to -1, we see that for $\rho \in [-20, 15]$ dB, the difference values are negative, indicating that OMA outperforms NOMA in this range. But when $\rho > 15$ dB, the values are positive, i.e., NOMA offers better link-layer rates. However, the particular ranges depend not only on the delay exponents but also on the power allocation coefficients. By increasing the transmission power of the weak user and reducing the transmission power of the strong user, we notice that the range is reduced. That range expands when we do the inverse. Also, when the delay becomes more stringent, e.g., $\beta_1 = \beta_2 = -2$, the zero crossing moves from 15 to 25 dB.

Figure 6 shows the difference of the EC in NOMA and the EC in OMA for the strong user. This curve starts initially

at zero, then increases to a certain maximum and starts decreasing without bound at high values of the transmit SNR. This confirms Lemma 3. We note that the maximum of these curves decreases when the delay becomes more stringent.

To investigate the impact of ρ on the performance of the total link-layer rate for the two-user system, in Fig. 7 the plots for V_N in NOMA and V_O in OMA, versus the transmit SNR are depicted for various delay exponents. The curves demonstrate that for both NOMA and OMA, the total EC for the two users starts at the initial value of 0 and then increases with the transmit SNR, as outlined in Lemma 4. When ρ is very small, the total link-layer rate for the two user in NOMA, V_N, increases faster than V_O in OMA. On the contrary, with the increase of the transmit SNR, V_O becomes gradually higher than V_N. At very high values of the transmit SNR, the gap
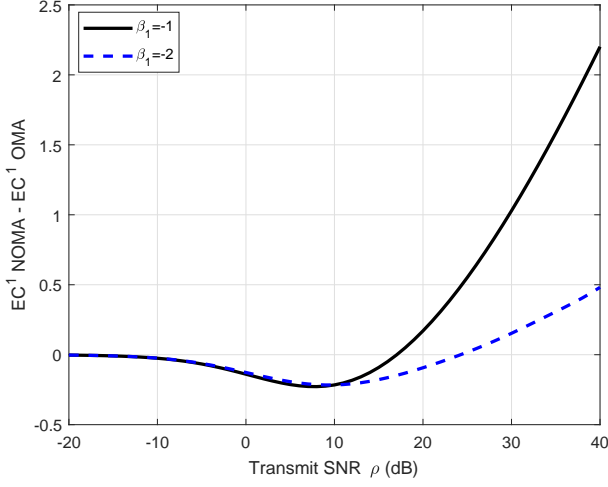
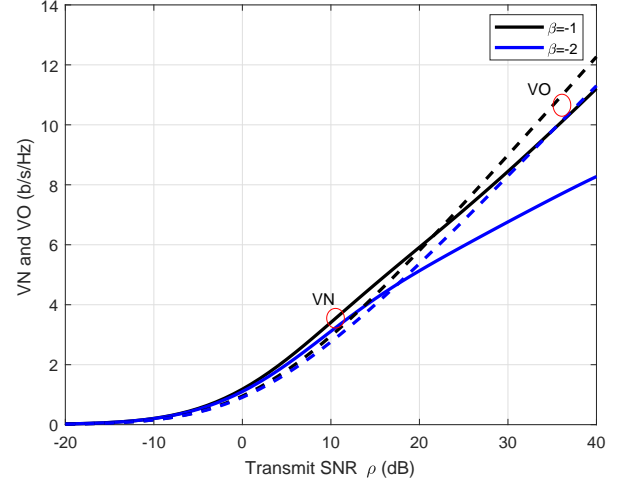Fig. 5. $E_c^1 - \tilde{E}_c^1$ versus ρ, for several values of the normalized delay exponent.

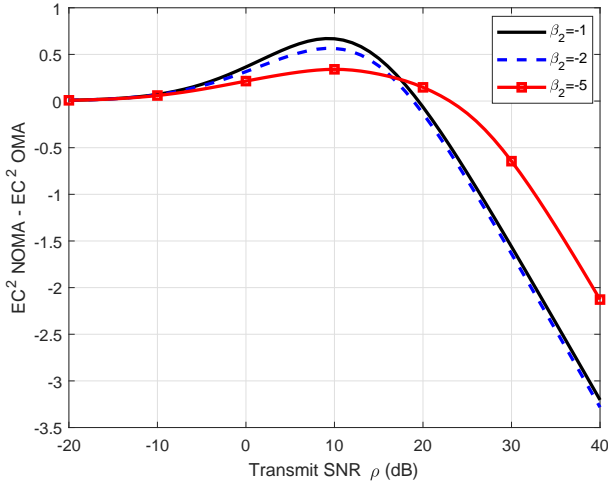Fig. 7. V_N and V_O versus ρ, for several values of normalized delay exponent.

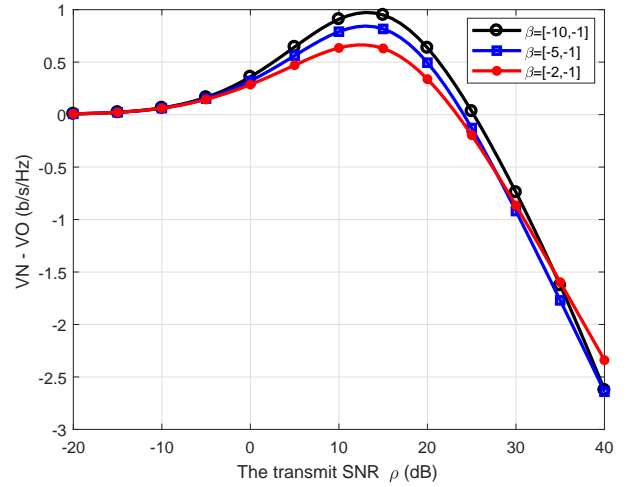Fig. 6. $E_c^2 - \tilde{E}_c^2$ versus ρ, for various normalized delay exponent.

Fig. 8. V_N - V_O versus ρ for various normalized delay.

between the sum EC with NOMA and OMA increases further. Finally, when the delay becomes more stringent, the sum EC of both NOMA and OMA decreases.

Finally, Figs 8 and 9 depict the sum ECs versus ρ, for several values of the (negative) normalized delay exponent. In Fig. 8, the delay of the strong user is fixed, while the delay exponent of the weak user varies. It is shown that in this case, the highest delay QoS (i.e., the smallest negative normalized delay exponent) of the weak user corresponds to the highest gap between the sum ECs $V_N - V_O$. On the other hand, when the delay of the weak user is fixed, Fig. 9 shows that the smallest delay Qos (i.e., the highest negative normalized delay exponent) for the strong user corresponds to the largest gap in $V_N - V_O$.

The curve of $V_N - V_O$ starts at zero, increases to a maximum, and returns to negative values. The transition to zero is at around $\rho = 25$, and $\rho = 20$ respectively for the figures 8 and 9. That means from 0 to 25dB (20dB in the Figure 9), the total link-layer rate of NOMA is higher than the OMA one. And when ρ becomes larger than this transition point, the total link-layer rate of OMA outperforms the NOMA one.

IV. CONCLUSIONS AND FUTURE WORK

The concept of the EC enabled us to study the achievable data-link layer rates when the delay QoS guarantees are in place in the form of delay exponents. We investigated the EC for the uplink of a two-user NOMA network, assuming a Rayleigh block fading channel. We derived novel closed-form expressions for the ECs of the two users and provided a comparison between NOMA and OMA. In NOMA networks, we showed that the ECs of both users decrease as the delay
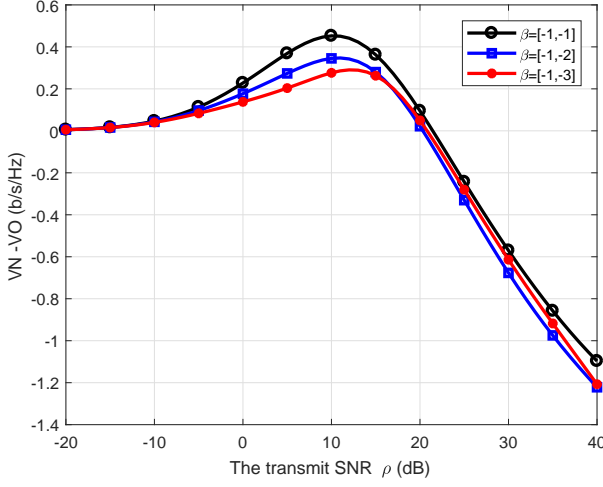
Fig. 9. V_N - V_O versus ρ for various normalized delay.

constraints become stringent. On the other hand, at high transmit SNRs, the EC of the weak user can surpass the EC of the strong user, as the latter is limited due to interference. This provides the possibility of switching between NOMA and OMA according to the individual users' delay constraints and transmit powers. In future work, the impact of user pairing will also be investigated.

APPENDIX I

For the second user, we have:

$$
\begin{aligned}
E_c^2 &= \frac{1}{\beta_2} \log_2 \left(2 \int_0^\infty (\frac{\rho P_2}{1 + \rho P_1 x_1})^{\beta_2} e^{-x_1} \right. \\
&\left. \times \int_{x_1}^\infty \left(\frac{1 + \rho P_1 x_1}{\rho P_2} + x_2 \right)^{\beta_2} e^{-x_2} dx_2 dx_1 \right).
\end{aligned}
$$

Set $z = \frac{1 + \rho P_1 x_1}{\rho P_2} + x_2$, which means we have $x_2 = z - \frac{1 + \rho P_1 x_1}{\rho P_2}$ and $dx_2 = dz$. Then,

$$
\begin{aligned}
E_c^2 &= \frac{1}{\beta_2} \log_2 \left(2 e^{\frac{1}{\rho P_2}} \int_0^\infty (\frac{\rho P_2}{1 + \rho P_1 x_1})^{\beta_2} e^{-x_1} e^{\frac{P_1 x_1}{P_2}} \right. \\
&\left. \int_{\frac{1 + \rho x_1}{\rho P_2}}^\infty z^{\beta_2} e^{-z} dz dx_1 \right) \\
&= \frac{1}{\beta_2} \log_2 \left(2 (\rho P_2)^{\frac{\beta_2}{2}} e^{\frac{1}{2\rho P_2}} \int_0^\infty (1 + \rho P_1 x_1)^{-\beta_2} \right. \\
&(1 + \rho x_1)^{\frac{\beta_2}{2}} e^{\frac{(2P_1 - 2P_2 - 1)x_1}{2P_2}} \left[\mathbf{W}_{\frac{\beta_2}{2}, \frac{1 + \beta_2}{2}} (\frac{1 + \rho x_1}{\rho P_2}) \right] dx_1 \right) \\
&= \frac{1}{\beta_2} \log_2 \left(2 P_2 (\rho P_2)^{\beta_2} e^{\frac{1}{\rho P_2}} e^{-\frac{(P_1 - P_2)}{\rho P_2}} \right. \\
&\left. \int_{\frac{1}{\rho P_2}}^\infty P_2^{-\beta_2} (1 + \rho P_1 y)^{-\beta_2} e^{(P_1 - P_2)y} \Gamma(1 + \beta_2, y) dy \right),
\end{aligned}
$$

where $\mathbf{W}$ is the Whittaker W function.

Using the binomial expansion, we have $(1 + \rho P_1 y)^{-\beta_2} = \sum_{j=0}^{-\beta_2} \binom{-\beta_2}{j} (\rho P_1 y)^j$. and we get the expression given in (9).

APPENDIX II

By inserting $\rho \to 0$ into (8) and (9), we get 1) of Lemma 1, i.e.,

$$
\lim_{\rho \to 0} (E_c^1 - \widetilde{E}_c^1) = \frac{1}{\beta_1} \log_2 \left(\frac{\mathbb{E}[(1 + \rho P_1 |h_1|^2)^{\beta_2}]}{\mathbb{E}[(1 + \rho |h_1|^2)^{\frac{\beta_2}{2}}]} \right) = 0,
$$

$$
\lim_{\rho \to 0} (E_c^2 - \widetilde{E}_c^2) = \frac{1}{\beta_2} \log_2 \left(\frac{\mathbb{E}[(1 + \frac{\rho P_2 |h_2|^2}{1 + \rho P_1 |h_1|^2})^{\beta_2}]}{\mathbb{E}[(1 + \rho |h_2|^2)^{\frac{\beta_2}{2}}]} \right) = 0.
$$

In the same way, by inserting $\rho \to \infty$ into (8) and (9), we get 2) in Lemma 1, given below.

$$
\lim_{\rho \to \infty} E_c^2 \to \frac{1}{\beta_2} \log_2 (\mathbb{E}[(1 + \frac{P_2 |h_2|^2}{P_1 |h_1|^2})^{\beta_2}]),
$$

$$
\lim_{\rho \to \infty} (E_c^1 - \widetilde{E}_c^1) = \frac{1}{\beta_1} \log_2 (\rho^{\frac{\beta_1}{2}} \frac{\mathbb{E}[(\frac{1}{\rho} + P_1 |h_1|^2)^{\beta_2}]}{\mathbb{E}[(\frac{1}{\rho} + |h_1|^2)^{\frac{\beta_2}{2}}]}) = \infty,
$$

$$
\lim_{\rho \to \infty} (E_c^2 - \widetilde{E}_c^2) = \frac{1}{\beta_2} \log_2 \left(\frac{\mathbb{E}[(\frac{\frac{1}{\rho} + P_1 |h_1|^2 + P_2 |h_2|^2}{\frac{1}{\rho} + P_1 |h_1|^2})^{\beta_2}]}{\rho^{\frac{\beta_2}{2}} \mathbb{E}[(\frac{1}{\rho} + |h_2|^2)^{\frac{\beta_2}{2}}]} \right) = -\infty.
$$

APPENDIX III

To analyze the trends of E_c^1 and $\widetilde{E}_c^1$ with respect to ρ, we start with

$$
\begin{aligned}
\frac{\partial E_c^1}{\partial \rho} &= \frac{1}{\beta_1 \ln 2} \frac{\left(\mathbb{E}[(1 + \rho P_1 |h_1|^2)^{\beta_1}] \right)'}{\mathbb{E}[(1 + \rho P_1 |h_1|^2)^{\beta_1}]} \\
&= \frac{P_1}{\ln 2} \frac{\mathbb{E}[|h_1|^2 (1 + \rho P_1 |h_1|^2)^{\beta_1 - 1}]}{\mathbb{E}[(1 + \rho P_1 |h_1|^2)^{\beta_1}]} \geq 0.
\end{aligned}
$$

Similarly, for user 1 in OMA we have

$$
\begin{aligned}
\frac{\partial \widetilde{E}_c^1}{\partial \rho} &= \frac{1}{\beta_1 \ln 2} \frac{\left(\mathbb{E}[(1 + \rho |h_1|^2)^{\frac{\beta_1}{2}}] \right)'}{\mathbb{E}[(1 + \rho |h_1|^2)^{\frac{\beta_1}{2}}]} \\
&= \frac{1}{2 \ln 2} \frac{\mathbb{E}[|h_1|^2 (1 + \rho |h_1|^2)^{\frac{\beta_1}{2} - 1}]}{\mathbb{E}[(1 + \rho |h_1|^2)^{\frac{\beta_1}{2}}]} \geq 0.
\end{aligned}
$$

Then, we get that

$$
\begin{aligned}
\frac{\partial (E_c^1 - \widetilde{E}_c^1)}{\partial \rho} &= \frac{P_1}{\ln 2} \frac{\mathbb{E}[|h_1|^2 (1 + \rho P_1 |h_1|^2)^{\beta_1 - 1}]}{\mathbb{E}[(1 + \rho P_1 |h_1|^2)^{\beta_1}]} \\
&- \frac{1}{2 \ln 2} \frac{\mathbb{E}[|h_1|^2 (1 + \rho |h_1|^2)^{\frac{\beta_1}{2} - 1}]}{\mathbb{E}[(1 + \rho |h_1|^2)^{\frac{\beta_1}{2}}]},
\end{aligned}
$$

and $\lim_{\rho \to 0} (\frac{\partial (E_c^1 - \widetilde{E}_c^1)}{\partial \rho}) = \frac{(P_1 - \frac{1}{2})}{\ln 2} \mathbb{E}[|h_1|^2] \leq 0$. When $\rho >> 1$, we have

$$
\begin{aligned}
\frac{\partial (E_c^1 - \widetilde{E}_c^1)}{\partial \rho}) &= \frac{P_1}{\ln 2} \frac{\mathbb{E}[|h_1|^2 (\rho P_1 |h_1|^2)^{\beta_1 - 1}]}{\mathbb{E}[(\rho P_1 |h_1|^2)^{\beta_1}]} \\
&- \frac{1}{2 \ln 2} \frac{\mathbb{E}[|h_1|^2 (\rho |h_1|^2)^{\frac{\beta_1}{2} - 1}]}{\mathbb{E}[(\rho |h_1|^2)^{\frac{\beta_1}{2}}]} = \frac{1}{2 \rho \ln 2} \geq 0.
\end{aligned}
$$

When $\rho \to \infty$, this term approaches 0.

$E_c^2 = \frac{1}{\beta_2}\log_2(\mathbb{E}[(1 + \frac{\rho P_2|h_2|^2}{1+\rho P_1|h_1|^2})^{\beta_2}])$, and

$$\frac{\partial E_c^2}{\partial \rho} = \frac{1}{\beta_2 \ln 2}\frac{\left(\mathbb{E}[(1 + \frac{\rho P_2|h_2|^2}{1+\rho P_1|h_1|^2})^{\beta_2}]\right)'}{\mathbb{E}[(1 + \frac{\rho P_2|h_2|^2}{1+\rho P_1|h_1|^2})^{\beta_2}]}$$

$$= \frac{1}{\ln 2}\frac{\mathbb{E}[\frac{P_2|h_2|^2}{(1+\rho P_1|h_1|^2)^2}(1 + \frac{\rho P_2|h_2|^2}{1+\rho P_1|h_1|^2})^{\beta_2-1}]}{\mathbb{E}[(1 + \frac{\rho P_2|h_2|^2}{1+\rho P_1|h_1|^2})^{\beta_2}]} \geq 0.$$

In the same way, for the user 2 in OMA, we have

$$\frac{\partial \widetilde{E}_c^2}{\partial \rho} = \frac{1}{\beta_2 \ln 2}\frac{\left(\mathbb{E}[(1 + \rho|h_2|^2)^{\frac{\beta_2}{2}}]\right)'}{\mathbb{E}[(1 + \rho|h_2|^2)^{\frac{\beta_2}{2}}]}$$

$$= \frac{1}{2\ln 2}\frac{\mathbb{E}[|h_2|^2(1 + \rho|h_2|^2)^{\frac{\beta_2}{2}-1}]}{\mathbb{E}[(1 + \rho|h_2|^2)^{\frac{\beta_2}{2}}]} \geq 0, \text{ and}$$

$$\frac{\partial(E_c^2 - \widetilde{E}_c^2)}{\partial \rho} = \frac{1}{\ln 2}\frac{\mathbb{E}[\frac{P_2|h_2|^2}{(1+\rho P_1|h_1|^2)^2}(1 + \frac{\rho P_2|h_2|^2}{1+\rho P_1|h_1|^2})^{\beta_2-1}]}{\mathbb{E}[(1 + \frac{\rho P_2|h_2|^2}{1+\rho P_1|h_1|^2})^{\beta_2}]}$$

$$- \frac{1}{2\ln 2}\frac{\mathbb{E}[|h_2|^2(1 + \rho|h_2|^2)^{\frac{\beta_2}{2}-1}]}{\mathbb{E}[(1 + \rho|h_2|^2)^{\frac{\beta_2}{2}}]}.$$

When $\rho \to 0$, we have that $\lim_{\rho \to 0}(\frac{\partial(E_c^2 - \widetilde{E}_c^2)}{\partial \rho}) = \frac{(P_2 - \frac{1}{2})}{\ln 2}\mathbb{E}[|h_2|^2]$. When ρ is very large,

$$\frac{\partial(E_c^2 - \widetilde{E}_c^2)}{\partial \rho} = \frac{\mathbb{E}[\frac{P_2|h_2|^2}{\rho^2(\frac{1}{\rho}+P_1|h_1|^2)^2}(1 + \frac{\rho}{\rho}\frac{(P_2|h_2|^2)}{(\frac{1}{\rho}+P_1|h_1|^2)})^{\beta_2-1}]}{\ln 2\mathbb{E}[(1 + \frac{\rho}{\rho}\frac{P_2|h_2|^2}{(\frac{1}{\rho}+P_1|h_1|^2)})^{\beta_2}]}$$

$$- \frac{1}{2\ln 2}\frac{1}{\rho}\frac{\mathbb{E}[|h_2|^2(\frac{1}{\rho} + |h_2|^2)^{\frac{\beta_2}{2}-1}]}{\mathbb{E}[(\frac{1}{\rho} + |h_2|^2)^{\frac{\beta_2}{2}}]}$$

$$= \frac{\mathbb{E}[\frac{P_2|h_2|^2}{\rho^2(P_1|h_1|^2)^2}(1 + \frac{P_2|h_2|^2}{P_1|h_1|^2})^{\beta_2-1}]}{\ln 2\mathbb{E}[(1 + \frac{P_2|h_2|^2}{P_1|h_1|^2})^{\beta_2}]} - \frac{1}{2\ln 2}\frac{1}{\rho}\frac{\mathbb{E}[(|h_2|^2)^{\frac{\beta_2}{2}}]}{\mathbb{E}[(|h_2|^2)^{\frac{\beta_2}{2}}]}$$

$$= \frac{P_2}{\rho^2 P_1^2}\frac{\mathbb{E}[\frac{|h_2|^2}{(|h_1|^2)^2}(1 + \frac{P_2|h_2|^2}{P_1|h_1|^2})^{\beta_2-1}]}{\ln 2\mathbb{E}[(1 + \frac{P_2|h_2|^2}{P_1|h_1|^2})^{\beta_2}]} - \frac{1}{2\ln 2}\frac{1}{\rho}$$

$$= \frac{\frac{P_2}{P_1^2 \ln 2}A - \frac{1}{2\ln 2}\rho}{\rho^2},$$

where $A = \frac{\mathbb{E}[\frac{|h_2|^2}{(|h_1|^2)^2}(1 + \frac{P_2|h_2|^2}{P_1|h_1|^2})^{\beta_2-1}]}{\mathbb{E}[(1 + \frac{P_2|h_2|^2}{P_1|h_1|^2})^{\beta_2}]}$, unrelated to ρ. Hence, when ρ is very large, $\frac{\partial(E_c^2 - \widetilde{E}_c^2)}{\partial \rho}$ can be approximated by $-\frac{1}{2\ln 2}\frac{1}{\rho}$, and it gradually approaches 0 when $\rho \to \infty$.

APPENDIX V

Note that $V_N = E_c^1 + E_c^2$. By using Lemma 1, we have $\lim_{\rho \to 0}(V_N) = 0$ and $\lim_{\rho \to \infty}(V_N) = \infty$. Then, we get that

$$\frac{\partial V_N}{\partial \rho} = \frac{\partial(E_c^1 + E_c^2)}{\partial \rho} = \frac{P_1}{\ln 2}\frac{\mathbb{E}[|h_1|^2(1 + \rho P_1|h_1|^2)^{\beta_1-1}]}{\mathbb{E}[(1 + \rho P_1|h_1|^2)^{\beta_1}]}$$

$$+ \frac{1}{\ln 2}\frac{\mathbb{E}[\frac{P_2|h_2|^2}{(1+\rho P_1|h_1|^2)^2}(1 + \frac{\rho P_2|h_2|^2}{1+\rho P_1|h_1|^2})^{\beta_2-1}]}{\mathbb{E}[(1 + \frac{\rho P_2|h_2|^2}{1+\rho P_1|h_1|^2})^{\beta_2}]} \geq 0.$$

When $\rho \to 0$, we have $\lim_{\rho \to 0}(\frac{\partial V_N}{\partial \rho}) = \frac{P_1}{\ln 2}\mathbb{E}[|h_1|^2] + \frac{P_2}{\ln 2}\mathbb{E}[|h_2|^2]$. When $\rho \to \infty$, we get that

$$\lim_{\rho \to \infty}\frac{\partial V_N}{\partial \rho} = \frac{1}{\rho \ln 2} + \frac{\mathbb{E}[\frac{P_2|h_2|^2}{(P_1|h_1|^2)^2}(1 + \frac{P_2|h_2|^2}{P_1|h_1|^2})^{\beta_2-1}]}{\rho^2 \ln 2\mathbb{E}[(1 + \frac{P_2|h_2|^2}{P_1|h_1|^2})^{\beta_2}]} = 0.$$

For V_O in the case of OMA, we note that $V_O = \widetilde{E}_c^1 + \widetilde{E}_c^2$. By using Lemma 1, we have $\lim_{\rho \to 0}(V_0) = 0$ and $\lim_{\rho \to \infty}(V_0) = \infty$. Then,

$$\frac{\partial V_0}{\partial \rho} = \frac{\partial(\widetilde{E}_c^1 + \widetilde{E}_c^2)}{\partial \rho} = \frac{1}{2\ln 2}\frac{\mathbb{E}[|h_1|^2(1 + \rho|h_1|^2)^{\frac{\beta_1}{2}-1}]}{\mathbb{E}[(1 + \rho|h_1|^2)^{\frac{\beta_1}{2}}]}$$

$$+ \frac{1}{2\ln 2}\frac{\mathbb{E}[|h_2|^2(1 + \rho|h_2|^2)^{\frac{\beta_2}{2}-1}]}{\mathbb{E}[(1 + \rho|h_2|^2)^{\frac{\beta_2}{2}}]} \geq 0.$$

When $\rho \to 0$, we have $\lim_{\rho \to 0}(\frac{\partial V_O}{\partial \rho}) = \frac{1}{2\ln 2}\mathbb{E}[|h_1|^2] + \frac{1}{2\ln 2}\mathbb{E}[|h_2|^2]$. When $\rho \to \infty$, we have that $\lim_{\rho \to \infty}(\frac{\partial V_O}{\partial \rho}) = \lim_{\rho \to \infty}(\frac{1}{2\rho \ln 2} + \frac{1}{2\rho \ln 2}) = \lim_{\rho \to \infty}(\frac{1}{\rho \ln 2})$, which equals to 0.

REFERENCES

[1] S. R. Islam, , N. Avazov, O. A. Dobre, and K.-S. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges," *IEEE Commun. Surv, & Tut.*, vol. 19, no. 2, pp. 721–742, 2016.

[2] Y. Kanaras and I. D. A. Chorti, M. Rodrigues, "An optimum detection for a spectrally efficient non orthogonal FDM system," in *Proc. 13th Int. OFDM WS*, Aug 2008, pp. 65–68.

[3] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. IEEE VTC Spring*, 2013, pp. 1–5.

[4] W. Yu, L. Musavian, and Q. Ni, "Link-layer capacity of NOMA under statistical delay QoS guarantees," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4907–4922, 2018.

[5] ——, "Tradeoff analysis and joint optimization of link-layer energy efficiency and effective capacity toward green communications," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3339–3353, 2016.

[6] D. Wu and R. Negi, "Effective capacity: a wireless link model for support of quality of service," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 630–643, 2003.

[7] J. Tang and X. Zhang, "Cross-layer modeling for quality of service guarantees over wireless links," *IEEE Trans. Wireless Commun.*, vol. 6, no. 12, pp. 4504–4512, 2007.

[8] A. Chorti and H. V. Poor, "Achievable secrecy rates in physical layer secure systems with a helping interferer," in *2012 Int Conf. Computing, Networking Commun. (ICNC)*, Maui, HI, 2012, pp. 18–22.

[9] W. Yu, A. Chorti, L. Musavian, H. V. Poor, and Q. Ni, "Effective secrecy rate for a downlink NOMA network," *IEEE Trans. Wireless Commun.*, pp. 5673–5690, Dec. 2019.

[10] 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 Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7244–7257, 2016.

[11] N. Zhang, J. Wang, G. Kang, and Y. Liu, "Uplink nonorthogonal multiple access in 5G systems," *IEEE Commun. L.*, vol. 20, no. 3, pp. 458–461, 2016.

[12] L. Fan, S. Jin, C.-K. Wen, and H. Zhang, "Uplink achievable rate for massive MIMO systems with low-resolution ADC," *IEEE Commun. L.*, vol. 19, no. 12, pp. 2186–2189, 2015.

[13] H.-C. Yang and M.-S. Alouini, *Order statistics in wireless communications: diversity, adaptation, and scheduling in MIMO and OFDM systems.* Cambridge University Press, 2011.