Performance Analysis of Uplink NOMA-Relevant Strategy Under Statistical Delay QoS Constraints
Mylène Pischella, Arsenia Chorti, Inbar Fijalkow

To cite this version:

HAL Id: hal-02553369
https://hal-cnam.archives-ouvertes.fr/hal-02553369
Submitted on 15 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Performance Analysis of Uplink NOMA-Relevant Strategy Under Statistical Delay QoS Constraints

Mylene Pischella, Senior Member, IEEE, Arsenia Chorti, Member, IEEE and Inbar Fijalkow, Senior Member, IEEE

Abstract—A new multiple access (MA) strategy, referred to as non orthogonal multiple access - Relevant (NOMA-R), allows selecting NOMA when this increases all individual rates, i.e., it is beneficial for both strong(er) and weak(er) individual users. This letter provides a performance analysis of the NOMA-R strategy in uplink networks with statistical delay constraints. Closed-form expressions of the effective capacity (EC) are provided in two-users networks, showing that the strong user always achieves a higher EC with NOMA-R. Regarding the networks sum EC, there are distinctive gains with NOMA-R, particularly under stringent delay constraints.

Index Terms: NOMA, effective capacity, QoS delay constraint.

I. INTRODUCTION

Due to the strict delay requirements of many emerging ultra reliable low latency communication (URLLC) applications in beyond fifth generation (B5G) networks, the investigation of the interplay between statistical delay quality of service (QoS) constraints and wireless propagation conditions is highly timely. In this context, employing the link layer service (QoS) constraints and wireless propagation conditions is the first attempt to propose an adaptive MA strategy, called NOMA-Relevant (NOMA-R). In NOMA-R, clusters of users employ NOMA only when it is beneficial for all of them in terms of their individual rates.

This letter is the first EC performance analysis of adaptive MA strategies. Its contributions are the following: (i) we evaluate the probability of using NOMA when NOMA-R is employed; (ii) we use the EC as the performance metric, provide new analytic expressions of the EC with NOMA-R and compare them to those derived in the two-users case for NOMA and OMA in [8]; (iii) we prove that NOMA-R is the strategy that maximizes the EC of the strong user, whereas it always outperforms OMA but not NOMA for the weak user; (iv) with respect to this latter aspect, this loss in EC for the weak user becomes negligible under stringent delay constraints; (v) numerical results also show that this conclusion holds for a larger number of users. This letter consequently proves that NOMA-R is a very efficient strategy for delay constrained applications.

II. PROBABILITY OF USING NOMA WITH NOMA-R

A. System model

Let us consider a network with \( K \) users employing either OMA, NOMA or NOMA-R in the uplink. The achievable rate of user \( k \in S_K = \{1, ..., K\} \) is denoted in the following by \( R_k, \hat{R}_k \) and \( \tilde{R}_k \) for NOMA, OMA and NOMA-R, respectively. We assume that the independent and identically distributed (i.i.d.) fading channel coefficients in the links to the base station (BS), denoted by \( h_k \), follow unit variance Rayleigh distributions. The channel gains \( x_k = |h_k|^2 \) are assumed ordered in decreasing order, so that \( x_k \leq x_{k+1} \forall k \in \{1, ..., K-1\} \), and, their distributions can be found by using the theory of order statistics [10]. We denote by \( \rho = \frac{1}{P_k} \) the transmit SNR with \( N \) the additive white Gaussian noise power in each link (assumed the same for all links for simplicity). The transmit power used by user \( k \) is denoted by \( P_k \) so that its received SNR is \( \rho P_k, k \in S_K \). Let us assume that a cluster \( S \) of users in \( S_K \) with cardinality \( |S| \) is chosen for NOMA. The achievable rates (in bits/s/Hz) of the \( k \)th user in \( S \), assuming perfect successive interference cancellation (SIC) decoding, can be expressed as:

\[
R_k = \frac{|S|}{K} \log_2 \left( 1 + \frac{\rho P_k x_k}{1 + \sum_{j \in S, j < k} \rho P_j x_j} \right), \forall k \in S. \tag{1}
\]

On the other hand, for the \( k \)th user in \( S_K \setminus S \) the achievable rates with OMA are given as

\[
\hat{R}_k = \frac{1}{K} \log_2 (1 + \rho P_k x_k), \forall k \in S_K \setminus S. \tag{2}
\]

The coefficients \( |S|/K \) and \( 1/K \) account for fair division of the resources between the users employing NOMA and OMA.

1We note that [9] focused on proportional fairness.
respectively. Although in a standard NOMA network all users will employ NOMA, i.e., \( S \) and \( S_K \) coincide, in NOMA-R, \( S \) is a subset of \( S_K \). The formalization of the NOMA-R criterion stems from the requirement that \( S \) includes all users whose achievable rates are greater with NOMA than with OMA, i.e.,

\[
1 + \frac{\rho P_k x_k}{1 + \sum_{j \in S} \rho P_j x_j} \geq (1 + \rho P_k x_k)^{\frac{1}{\beta_k}}.
\]

Users in \( S_K \setminus S \) employ OMA. Consequently, the data rate in NOMA-R is \( \tilde{R}_k = R_k \forall k \in S \) and \( \tilde{R}_k = \hat{R}_k \forall k \notin S \). In terms of implementation, the BS identifies the subsets \( S \) using its knowledge of the network’s full channel state information (CSI). It then transmits to each user a one-bit feedback indicating whether they should use OMA or NOMA and either the user’s index in the OMA subset or the NOMA cluster index if several disjoint NOMA clusters are selected. Therefore, NOMA-R imposes at most one-bit signalling overhead with respect to OMA.

The EC in bits/s/Hz of user \( k \in S_K \) is defined as [1]:

\[
E_{c_k} = -\frac{1}{\theta_k T_j B} \ln \left( \mathbb{E}[e^{-\theta_k T_j B r_k}] \right) = \frac{1}{\beta_k} \log_2 \left( \mathbb{E}[e^{\beta_k \ln(2) r_k}] \right)
\]

where \( r_k \) is the achievable rate of user \( k \) (equal to \( R_k, \tilde{R}_k, \) or \( \hat{R}_k \) if the user employs NOMA, OMA or NOMA-R, respectively), \( T_j \) is the symbol period and \( B \) is the occupied bandwidth, \( \theta_k \), known as the QoS exponent [1], is the exponent of the exponential decay of the buffer overflow probability. Under a constant packet arrival rate assumption, the EC is defined as the maximum achievable rate such that a target delay-bound violation probability is met. The more stringent the delay requirement, the larger the delay exponent \( \theta_k \). To simplify the notation, we define \( \beta_k = -\frac{\theta_k T_j B}{\ln(2)} \) as the negative QoS exponent. Closed form expressions of the EC when OMA and NOMA are employed were derived in [8] for \( K = 2 \) and are not repeated in the present for compactness.

B. Probability of using NOMA while in NOMA-R for \( K \geq 2 \)

When NOMA-R selects NOMA for \( k \in S \), the sum rate can easily be shown to be equal to \( \frac{1}{S} \sum_{k \in S} \log_2 \left( 1 + \sum_{k \in S} \rho P_k x_k \right) \). Consequently, the subset \( S \) of \( S_K \) that maximizes the sum rate while satisfying (3) is selected by NOMA-R. Several disjoint clusters may also be selected if they independently verify (3).

The probability of using NOMA when NOMA-R is employed, denoted by \( \tau_K \) in the following, is the union of the probabilities to verify (3) for any subset \( S \subseteq S_K \) with \( |S| \geq 2 \) and its analytical derivation is very evolved. As an illustrative example, let us assume \( S = \{1, \ldots, k\} \) with \( k \leq K \). Let \( y_k = \rho P_k x_k \) be the weighted kth order statistics and \( z_k = \sum_{i=1}^{k-1} \rho P_i x_i \) the weighted sum of the lowest \((k-1)\)th order statistics. Then \( \tau_k \) is equal to:

\[
\tau_k = P_r \left( \bigcap_{i=2:k} \left( \frac{(1 + y_i) - (1 + y_i)^{\frac{1}{\beta_i}}}{(1 + y_i)^{\frac{1}{\beta_i}} - 1} \geq z_i \right) \right).
\]

For the specific case \( k = K \), the joint probability density function (pdf) of \((y_K, z_K)\) can be derived by using the moment generating function (MGF) of the weighted sum of the lowest order statistics, denoted as \( M_{z_K} \), as in [10]. The pdf of \( z_k \) can be obtained by using the following properties:

\[
M_{z_K} = \prod_{j=1}^{K-1} M_{z_j}(\rho P_j x_j) \quad \text{and} \quad L^{-1}\left(M_{z_j}(\rho P_j x_j)\right)(t) = \frac{1}{\rho P_j x_j} \left( e^{t \rho P_j x_j} - 1 \right),
\]

where \( L^{-1} \) is the inverse Lagrange transform. Consequently, the joint pdf \( f_{z_K, y_K} = \tilde{g} \) can be derived from that of \( z_k \) and \( y_K \) in [10] eq. (3.41). However, a closed-form expression of \( Pr \left( \frac{(1+y_K) - (1+y_K)^{\frac{1}{\beta}}}{(1+y_K)^{\frac{1}{\beta}} - 1} \geq z_K \right) \) cannot be obtained, and similarly to the conclusion in [11] Section VD, it should be calculated with a mathematical software. Moreover when \( k < K \), to the best of our knowledge, the joint pdf of \((y_K, z_K)\) is yet unknown. Consequently, when \( K > 2 \), (4) cannot be evaluated analytically with reasonable effort. For all these reasons, our analytical study is limited to \( K = 2 \) while we provide numerical results for \( K > 2 \).

Finally, examining the case non i.i.d. channel coefficients, we note that the case of non-identical exponential distributions can be treated as in [12], while the case of non independent coefficients as in [13]. If full CSI is not available at the BS, CSI uncertainties can be inserted in users’ distributions to derive an MA selection strategy [13], and, when \( K = 2 \), (3) can be formulated as a binary hypothesis testing problem [14].

C. Probability of using NOMA while in NOMA-R for \( K = 2 \)

In the following, we consider the two-users case and call user 2 the strong user, and user 1 the weak user. As \( R_1 \geq \tilde{R}_1 \) is always fulfilled, the NOMA-R strategy is used whenever \( R_2 \geq \tilde{R}_2 \). The NOMA-R condition consequently simplifies to \( x_2 \geq \rho^2 x_1^2 P_1^2 - 1 \) and \( \tau_2 = \tau(\rho) = Pr \left( x_2 \geq \rho^2 x_1^2 P_1^2 - 1 \right) \). Using the theory of order statistics, the pdf of \( x_1 \) is \( 2 e^{-x_1} \), the pdf of \( x_2 \) is \( 2 e^{-x_2} \left(1 - e^{-x_2}\right)\) and the joint pdf of \((x_1, x_2)\) is \( 2 e^{-x_1} e^{-x_2} \). Then \( \tau(\rho) \) is equal to:

\[
\tau = \int_{x_1=0}^{\infty} \int_{x_2=x_1}^{\infty} 2 e^{-x_1} e^{-x_2} dx_2 dx_1 + \int_{x_1=0}^{\infty} \int_{x_2=0}^{x_1} 2 e^{-x_1} e^{-x_2} dx_2 dx_1
\]

\[
= f(\rho) + g(\rho)
\]

(5)

where

\[
f(\rho) = 1 - e^{-\rho^2 P_2 + \sqrt{\rho^2 P_2^2 + 4 P_1^2}} P_1^2 P_2^2
\]

\[
g(\rho) = \sqrt{\rho^2 P_2^2 + 4 P_1^2} \left( 1 - \text{erf} \left( \frac{2 P_2 + \sqrt{P_2^2 + 4 P_1^2}}{2 P_2 P_1^2} \right) \right) \frac{P_1^2}{P_1^2 + P_2^2}
\]

(7)

and \( \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-t^2} dt \). The boundary in both integrals in (5) is due to the fact that \( x_2 \) should always be such that \( x_2 \geq x_1 \), but \( \rho^2 x_1^2 P_1^2 - 1 \) is lower than \( x_1 \) if \( x_1 \leq \frac{P_2 + \sqrt{P_2^2 + 4 P_1^2}}{2 P_2 P_1^2} \). \( \tau(\rho) \) is a monotonically decreasing function with respect to \( \rho \). (5) is validated with Monte-Carlo simulations, shown in Fig. 1 assuming that \( P_1 = P_2 = 1 \).
Lemma 2. \( \tau(\rho) \) tends to 1 when \( \rho \) tends to 0 and \( \tau(\rho) \) tends to 0 when \( \rho >> 1 \).

Proof. When \( \rho \to 0 \), \( f(\rho) \to 1 \). Moreover, \( \text{erf}(x) \approx 1 - e^{-x^2}/(x\sqrt{\pi}) \) when \( x \gg 1 \), which implies that \( g(\rho) \approx ae^{-b/\rho} \) with \((a, b)\) two strictly positive constants. Therefore, \( g(\rho) \to 0 \) when \( \rho \to 0 \). Furthermore, when \( \rho >> 1 \), \( f(\rho) \to 0 \). As \( \text{erf}(x) \approx \frac{2}{\sqrt{\pi}}x e^{-x^2} \) when \( x \to 0 \), \( g(\rho) \approx a_1 + b_1\frac{e^{2x^2}}{\rho} \) where \((a_1, a_2, b_1, b_2)\) are constants and \( g(\rho) \to 0 \).

III. NOMA-R EFFECTIVE RATES

The ECs of user \( k = 1, 2 \) when employing NOMA, OMA or NOMA-R are denoted by \( E^k_{c,N} \), \( E^k_{c,O} \) and \( E^k_{c,R} \), respectively.

A. NOMA-R EC of user 1

We hereafter derive closed-form expressions of the EC with NOMA-R. As \( \tau \) corresponds to the proportion of time spent in NOMA and \((1 - \tau)\) is the proportion of time spent in OMA, the EC of user \( k \in \{1, 2\} \) when using the NOMA-R strategy is employed is given as:

\[
E_{c,R}^k = \frac{1}{\beta_k} \log_2 \left( E \left[ e^{\beta_k R_k + \beta_k (1-\tau) R_k} \right] \right). \tag{8}
\]

For user 1, the EC achieved with the NOMA-R strategy is:

\[
E_{c,R}^1 = \frac{1}{\beta_1} \log_2 \left( E \left[ (1 + \rho P_1 x_1) \frac{\beta_1 (\tau + 1)}{2} \right] \right)
= \frac{1}{\beta_1} \log_2 \left( \frac{2}{\rho P_1} U \left( 1, 2 + \frac{\beta_1 (\tau + 1)}{2}, \frac{2}{\rho P_1} \right) \right) \tag{9}
\]

where \( U(a, b, z) = \int_0^\infty e^{-zt} t^{a-1} (1+t)^{b-a-1} dt \) denotes the confluent hypergeometric function.

Lemma 2. \( E_{c,R}^1 \) is monotonically increasing with \( \rho \).

Proof. Let us consider \( \rho_1 \) and \( \rho_2 \) such that \( \rho_1 < \rho_2 \). For any value of \( \beta_1 \), let us define: \( \beta_{1,a} = \frac{\beta_1 (1 + \tau R_1)}{\rho_1} \) and \( \beta_{1,b} = \frac{\beta_1 (1 + \tau R_2)}{\rho_2} \). Then from [8 eq.(11)] and [9, 11], \( E_{c,R}^1(\beta_1, \rho_1) = E_{c,N}^1(\beta_{1,a}, \rho_1) \) and \( E_{c,R}^1(\beta_1, \rho_2) = E_{c,N}^1(\beta_{1,b}, \rho_2) \). As \( \tau(\rho) \) is a decreasing function with respect to \( \rho \), and \( \beta \) are negative, \( \beta_{1,a} \leq \beta_{1,b} \). Moreover, \( E_{c,R}^1(\beta_1, \rho) \) is increasing both with respect to \( \beta \) and to \( \rho \) according to [8]. Consequently, \( E_{c,N}^1(\beta_{1,a}, \rho_1) \leq E_{c,N}^1(\beta_1, \rho_2) \leq E_{c,N}^1(\beta_{1,b}, \rho_2) \) and:

\[
E_{c,R}^1(\beta_1, \rho_1) \leq E_{c,R}^1(\beta_1, \rho_2) \quad \forall \rho_1 < \rho_2 \tag{10}
\]

B. NOMA-R EC of user 2

For user 2, the NOMA-R EC is given by

\[
E_{c,R}^2 = \frac{1}{\beta_2} \log_2 \left( E \left[ \left( 1 + \rho P_2 x_2 \right) \frac{\beta_2 \tau}{1 + \rho P_1 x_1} \right] ^{\beta_2 \tau} \left( 1 + \rho P_2 x_2 \right) ^{\frac{\beta_2 (1 - \tau)}{2}} \right) \tag{11}
\]

Lemma 3. When \( \rho \) tends to 0, the EC with the NOMA-R strategy becomes equivalent to that of NOMA given in [3].

Proof. When \( \rho \to 0 \), \( \tau(\rho) \to 1 \) and therefore \( 1 + \rho P_2 x_2 \to 1 \). Consequently, \( E_{c,R}^2 \) tends to \( E_{c,N}^2 \).

Lemma 4. When \( \rho >> 1 \), the EC with the NOMA-R strategy becomes equivalent to that of OMA and its closed-form expression is given by:

\[
E_{c,R}^2 \approx \frac{1}{\beta_2} \log_2 \left( \left( \frac{\beta_2}{2} + 1 \right) \left( \rho P_2 \right) \frac{\beta_2 \tau}{2 - \beta_2 \tau} \right) \tag{12}
\]

Proof. When \( \rho >> 1 \), \( 1 + \frac{\rho P_2 x_2}{1 + \rho P_1 x_1} \to 1 \) because \( \tau \to 0 \). Then using \((1+x)^a \approx x^a\), the EC of user 2 becomes:

\[
E_{c,R}^2 \approx \frac{1}{\beta_2} \log_2 \left( E \left[ (\rho P_2 x_2) \frac{\beta_2 \tau}{2} \right] \right)
\approx \frac{1}{\beta_2} \log_2 \left( \int_0^\infty 2 \rho P_2 x_2 \frac{\beta_2 \tau}{2} e^{-x_2 (1 - e^{-x_2})} dx_2 \right).
\]

The integral’s closed-form expression leads to [12].

Theorem 1. The EC of user 1 is always larger with NOMA than with NOMA-R, while OMA is the worst strategy in terms of EC. Moreover, the EC of user 2 is always larger with the NOMA-R strategy than with OMA or NOMA.

Proof. The NOMA-R instantaneous rate of user 1 is equal to \( \hat{R}_1 = \frac{(1 + \tau R_1)}{2} \log_2 (1 + \rho P_1 x_1) \), according to [9]. Therefore \( \hat{R}_1 \leq \tilde{R}_1 \leq R_1 \). Then as \( \beta_1 \) is negative, \( e^{\beta_1 R_1} \leq e^{\beta_1 \hat{R}_1} \leq e^{\beta_1 \tilde{R}_1} \), and \( E^{\beta_1 \tilde{R}_1} \leq E^{\beta_1 \hat{R}_1} \leq E^{\beta_1 \tilde{R}_1} \), so that

\[
E_{c,N}^1 \geq E_{c,R}^1 \geq E_{c,O}^1 \tag{13}
\]

The NOMA-R rate of user 2 is \( \tilde{R}_2 = \max \{ R_2, \tilde{R}_2 \} \). Following the same steps as for user 1, we conclude that

\[
E_{c,R}^2 \geq E_{c,N}^2 \quad \text{and} \quad E_{c,O}^2 \geq E_{c,R}^2 \tag{14}
\]

Remark: \( E_{c,R}^2 \) asymptotically tends to either \( E_{c,N}^2 \) or \( E_{c,O}^2 \), both of which are monotonically increasing with \( \rho \). Moreover, contrary to the NOMA strategy, the NOMA-R strategy does not lead to a saturation of the EC of the strong user because of [14] and because \( E_{c,O}^2 \) increases without bound with \( \rho \) [8].
**IV. Numerical Results**

Performance comparisons with respect to EC with the NOMA-R strategy are provided for $P = [0.2, 0.8]$, $P = [0.05, 0.15, 0.8]$ and $P = [0.01, 0.04, 0.15, 0.8]$ when $K = 2$, 3 and 4, respectively. Unless otherwise stated, $\beta_k = -2$, $\forall k$. Fig. 2 validates the results of Theorem 1. Fig. 3 shows that the largest sum EC is always achieved with NOMA-R whatever the value of $K$. Moreover, the sum EC with NOMA-R coincides with that of NOMA when the SNR is lower than a minimum value $\rho_{\text{min}}$ that increases with $K$, because constraint (3) becomes less stringent when $K$ increases. Fig. 4 shows the dependency of the EC on $\beta_1$ when the SNR is equal to 35 dB, $\beta_K = -2$ and $\beta_j = \beta_1$, $\forall j < K$. The sum EC is larger with NOMA-R, except when $\beta_1$ approaches 0. The NOMA-R strategy is consequently more favorable when the target delay-bound violation probabilities are more stringent, especially for weak users.

**V. Conclusions**

In this letter the EC performance of an adaptive MA strategy, NOMA-R, was studied both analytically for $K = 2$ users and numerically for larger values of $K$. It was shown that NOMA-R is an advantageous strategy for delay constrained applications in BSG, e.g., URLLC, particularly as the users’ delay-outage probability constraints become more stringent.

**References**


