Balanced-QoS power allocation schemes for NOMA in the finite blocklength regime

Maria Cecilia Fernández Montefiore[1,2], Gustavo González[1], F. J. Lopez-Martinez[2] and Fernando Gregorio[1]

[1] Instituto de Investigaciones en Ingeniería Eléctrica "Alfredo Desages" (IIIE), UNS-CONICET, Dpto. Ingeniería Eléctrica y de Computadoras, Universidad Nacional del Sur (UNS), Bahía Blanca, Argentina.
[2] Research Centre for Information and Communication Technologies (CITIC-UGR), Dept. Signal Theory, Networking and Communications, University of Granada, 18071, Granada, Spain.
cecilia.fernandezmontefiore@uns.edu.ar

Abstract. Non-orthogonal multiple Access (NOMA) technology may find its way into sixth-generation (6G) mobile networks thanks to its ability to support heterogeneous services and improve spectral efficiency. In this context, shorter packet usage is imperative for reducing latency and strengthening reliability. However, this approach results in the liability of necessitating metrics adaptation to accommodate the finite blocklength (FBL) regime. This paper explores power allocation optimization for low-latency applications in multi-user downlink NOMA systems with imperfect successive interference cancellation (SIC) and FBL coding. The optimization aims to maximize the effective sum rate while complying with balanced QoS constraints. Two optimization approaches are presented: one optimizing power allocation based on a modified rate equation, and the other based on achievable signal-to-noise ratio (SNR). We obtained substantial performance improvements in achieving a balanced QoS among users while maintaining system efficiency.

Keywords: 6G networks · QoS · Power allocation schemes · Low-latency communication · multi-user NOMA.

1 Introduction

The constant evolution of wireless applications requires maintaining and upgrading the Quality-of-Service (QoS) standards for the three key services of the fifth generation of mobile communications (5G): ultra-reliable and low-latency communications (uRLLC), massive Machine-Type Communication (mMTC), and enhanced Mobile Broadband (eMBB). Stringent requirements regarding delay and reliability [1], an increased number of infrequent communicating devices sharing the same bandwidth [2], and substantial enhancements in bandwidth and capacity [3] are some challenges presented by the new advancements. Therefore, the main goal of 6G wireless networks is to improve technical factors such as the number of connections, latency, reliability, and throughput [4].

URLLC is an essential component of 5G and 6G networks, enabling mission-critical applications like enhanced vehicle-to-everything (eV2X), e-health, tactile internet, and haptic communications. The nature of this service demands the use of short packets with finite blocklength (FBL) codes to reduce physical layer transmission delays. As a result, the traditional Shannon capacity, which assumes infinite blocklength for perfect reliability, is no longer suitable for characterizing the maximum achievable data transmission rate while considering the constraints of transmission latency and finite data-packet length [5].

The concept of achievable effective rate (or Effective Capacity -EC- for simplification) is commonly used to evaluate user performance in various settings. EC denotes the highest possible arrival rate that a network can sustain while a specific latency condition is met [6]. The effective capacity model describes wireless channels in terms of functions easily translated into link-level QoS measurements [7]. This channel paradigm has several advantages, including ease of implementation and conversion into a QoS requirement, as the channel may be described in terms of QoS metrics. Addressing the finite blocklength (FBL) coding, [8] derived the achievable effective rate.

Traditional multiple-access technology is not easily scalable, and therefore unsuitable for some of the use cases envisioned for 6G. Non-orthogonal multiple access (NOMA) enables simultaneous communication over non-orthogonal resources by multiplexing users in the power domain and de-multiplexing them on the receiver side [9] [10]. Because this access technique incorporates the possibility of stronger and weaker users for multiplexing, the percentage of power allocated to each user is a variable that may be optimized. In [11], we demonstrated that NOMA outperforms OMA in a two-user system with heterogeneous constraints.

NOMA implementation requires superposition coding (SC) at the transmitter and successive interference cancellation (SIC) at the receiver. According to [12] and [13], users with better channel conditions are SIC-performing, whereas only users with the poorest channel conditions do not employ SIC. When evaluating real-world environments, to assume perfect interference cancellation at the receiver may not be accurate, as interference significantly reduces detection performance when it cannot be eliminated [14] [15]. The imperfect SIC will cause a residual error for every SIC-performing user [16].

In [17] we extended our previous work by including two practical impairments associated with these scenarios: the error caused by the incomplete interference cancellation and the possibility of considering several users. Nevertheless of these changes, NOMA still outperforms OMA under reasonable estimation errors regarding homogeneous and heterogeneous latency requirements for a fixed power allocation set-up.

This paper aims to improve the power allocation scheme for low-latency applications in the multi-user downlink of NOMA systems with imperfect SIC in the finite blocklength scenario, aiming to provide a balanced QoS among users. Differently from [17], the power allocation coefficients are now optimized using two different approaches: (i) from a modified rate equation; and (ii) from achievable SNR. As we will later see, substantial performance improvement can be attained thanks to power allocation optimization.

The rest of this paper is organized as follows. Section II introduces the system model and the achievable rates for a N-user system. Section III presents the formulation for the optimization problem with the approaches mentioned before. We present the optimization results and compare the performance of the optimized coefficients with the outcomes of earlier works in Section IV. Finally, conclusions and future research directions are discussed in Section V.

2 System Model

Let us present an N-user network, on which a base station (BS) communicates with downlink single-antenna users over a fading channel. In our system model, the strongest user is referred to as user 1, and the weakest is therefore denoted as user N. The channel gains for each user are h_i, which verify $|h_1| > |h_k| > |h_N|$. The investigated setup is intended for low-latency applications. Hence, the packets transmitted by the BS must be short, so a finite blocklength (FBL) formulation will be applied.

The NOMA principle states that the transmitted signal strength of weak users must be larger than that of strong users. By sequentially subtracting the decoded messages of weaker users from the received signal, SIC User k decodes and cancels interference caused by users $k+1$ to N and treats users 1 to $k-1$ as extra noise.

$$y_k = \underbrace{h_k \sum_{i=1}^{k-1} \left(\sqrt{\alpha_i P} x_i\right)}_{\text{interference}} + \underbrace{h_k \sum_{i=k+1}^{N} \left(\sqrt{\alpha_i P} x_i\right)}_{\text{to be removed by SIC}} + \underbrace{h_k \sqrt{\alpha_k P} x_k}_{\text{desired signal}} + \underbrace{w_k}_{\text{noise}}, \quad (1)$$

where α_i is the fraction of power allocated to user i, P is the total transmitted power, and w_k corresponds to additive white Gaussian noise. The power allocated for each user verifies $\sum_{i=1}^{N} \alpha_i = 1$. Moreover, the condition that the users and consequently channel gains are ordered from the strongest to the weakest is also represented by the following ordering of the α_i coefficient: $\alpha_1 < \alpha_i < \alpha_N$.

Let us consider the error produced by imperfect interference cancellation to represent a more realistic scenario. Usually, SIC can have a faulty performance due to the significant fading of the channels or channel estimation errors. For simplicity, yet without loss of generality, we will assume that there is no error in decoding, and error propagation in the SIC canceller is not considered in our analysis. Then, after removing users $k+1$ to N, we obtain:

$$\tilde{y_k} = \underbrace{h_k \sqrt{\alpha_k P} x_k}_{\text{desired signal}} + \underbrace{(h_k - \hat{h}_k) \sum_{i=k+1}^{N} \left(\sqrt{\alpha_i P} x_i\right)}_{\text{residual after imperfect SIC}} + \underbrace{h_k \sum_{i=1}^{k-1} \left(\sqrt{\alpha_i P} x_i\right) + w_k}_{\text{noise}}, \quad (2)$$

where $\hat{h}_k$ represents the imperfect estimation of the channel.

As for the effect of imperfect CSI, the strongest and weakest users are the extreme cases of this scenario. On the one hand, user N perceives multi-user interference (MUI) as additional noise, therefore SIC is not being used. On the other, user 1 is the one who is most affected by incorrect SIC, because its weaker signal is retrieved from the remainder after decoding and subtracting the stronger signals from other users.

Considering the channel estimation error E, which verifies $h_k = \hat{h}_k + E$, where $E \sim CN(0, \sigma_E^2)$. Then, the received SNR at the user k is given by:

$$\mathrm{S}NR_k = \frac{\alpha_k \rho |h_k|^2}{\sum_{i=1}^{k-1} \alpha_i \rho |h_k|^2 + \sum_{i=k+1}^{N} \alpha_i \rho \sigma_E^2 + 1}, \tag{3}$$

where $\rho = \frac{P}{N_0 B}$ is the transmit SNR.

The achievable rates for user k in the FBL regime are a function of $\mathrm{S}NR_k$ (the SNR of user k), the block length n and the error probability ϵ, and can be written as follows:

$$r_k = \log_2(1 + \mathrm{SNR}_k) - \sqrt{\frac{V_k}{n}} Q^{-1}(\epsilon), \tag{4}$$

where $Q^{-1}(\cdot)$ is the inverse of the Gaussian Q-function. Additionally, the channel dispersion for user k is:

$$V_k = 1 - (1 + \mathrm{SNR}_k)^{-2}. \tag{5}$$

The maximum arrival rate for a particular delay requirement is represented by the effective capacity C_k, determined by the user's delay exponent λ_k. A higher λ_k value indicates a more restrictive delay requirement.

$$C_k = -\frac{1}{\lambda_k n} \ln \left(\mathbb{E} \left[\epsilon + (1 - \epsilon)(1 + \mathrm{SNR}_k)^{2\Upsilon_k} e^{\psi_k \sqrt{V_k}} \right] \right), \tag{6}$$

where ln denotes the natural logarithm, $\mathbb{E}[\cdot]$ is the expectation operator, $\psi_k = \lambda_k \sqrt{n} Q^{-1}(\epsilon)$, and $\Upsilon_k = -\frac{\lambda_k n}{2 \ln 2}$.

Analytical formulations for the effective rate metrics are not included for simplicity, however, some approximate results can be calculated using the approach in [12]. For the case of Rayleigh fading, EC metrics are evaluated using numerical integration and ordered statistics, and these results are validated through Monte Carlo simulations.

3 Problem Formulation

The primary goal of this study is to achieve a balanced power allocation. This pursuit is to increment the sum rate while maintaining the QoS constraints. For this analysis, a three-user scenario with imperfect (SIC) and finite blocklength was considered. For this purpose, two different approaches are considered, which are detailed in the sequel.

3.1 Approach A: Modified rate equation.

The objective is to optimize the power allocation coefficients to maximize the sum rate, while keeping the individual user rates balanced, i.e., maintaining an achievable rate for each user above a target (equal) minimum rate r^t. Because it is not possible to formulate the optimization problem to include this restriction using the rate equation (4), the first approach to solve this problem is to relax the FBL assumption *only* for such restriction. It is worth remembering that the optimization variables are the power allocation coefficients of users 1 and 2. This is because there are two degrees of freedom, given that the sum of all coefficients must equal 1.

For simplicity in notation, $\rho|h_k|^2 = \gamma_k$ and $\rho\sigma_E^2 = \beta_E$. The optimization problem can be formulated as follows, where SR is the sum rate:

$$
\begin{aligned}
\max_{\alpha_1,\alpha_2} \quad SR =& \log_2\left(1+\frac{\alpha_1\gamma_1}{(1-\alpha_1)\beta_E+1}\right)\\
&+\log_2\left(1+\frac{(1-\alpha_1)\gamma 2}{\alpha_1\gamma_2+(1-\alpha_1-\alpha_2)\beta_E+1}\right)\\
&+\log_2\left(1+\frac{(1-\alpha_1-\alpha_2)\gamma 3}{(\alpha_1+\alpha_2)\gamma_3+1}\right)\\
&-\sqrt{\frac{1-\left(1+\frac{\alpha_1\gamma_1}{(1-\alpha_1)\beta_E+1}\right)^{-2}}{n}}Q^{-1}(\epsilon)\\
&-\sqrt{\frac{1-\left(1+\frac{(1-\alpha_1)\gamma 2}{\alpha_1\gamma_2+(1-\alpha_1-\alpha_2)\beta_E+1}\right)^{-2}}{n}}Q^{-1}(\epsilon)\\
&-\sqrt{\frac{1-\left(1+\frac{(1-\alpha_1-\alpha_2)\gamma 3}{(\alpha_1+\alpha_2)\gamma_3+1}\right)^{-2}}{n}}Q^{-1}(\epsilon)
\end{aligned}
$$

$$
\begin{aligned}
\text{s.t.} \quad & \alpha_1 \leq 1\\
& \alpha_1 \geq 0\\
& \alpha_2 \leq 1\\
& \alpha_2 \geq 0\\
& \alpha_1+\alpha_2 \leq 1\\
& \alpha_1(-\gamma_1-\beta_E 2^{r^t}+\beta_E) \leq (-2^{r^t}+1)(\beta_E+1)\\
& \alpha_1(\gamma_2-\beta_E)(2^{r^t}-1)+\alpha_2(-\gamma_2-\beta_E(2^{r^t}-1)) \leq (-2^{r^t}+1)(\beta_E+1)\\
& \alpha_1 2^{r^t}\gamma_3+\alpha_2 2^{r^t}\gamma_3 \leq -2^{r^t}+1+\gamma_3
\end{aligned} \tag{7}
$$

3.2 Approach B. Achievable SNR.

While the earlier approach allows to formulate the optimization problem with a number of feasible restrictions, it fails to consider the limitations inherent to a system operating with a finite blocklength. This may result in an artificially inflated rate, achieved by opting for a lower power allocation coefficient. While this fictitious rate may satisfy the conditions set for the target rate, it does not accurately reflect a short-packet scenario.

To accurately depict this scenario, it was necessary to revise the constraints. Instead of balancing the per-user achievable rates, we decided to maintain the per-user SNR above a minimum (equal) target (S^t). With this approach, it was not necessary to modify the rate equation (eq. 4) to formulate the optimization problem. Hence, this can be formulated as follows:

$$
\begin{aligned}
\max_{\alpha_1,\alpha_2} \quad SR = & \log_2\left(1+\frac{\alpha_1\gamma_1}{(1-\alpha_1)\beta_E+1}\right) \\
& + \log_2\left(1+\frac{(1-\alpha_1)\gamma 2}{\alpha_1\gamma_2+(1-\alpha_1-\alpha_2)\beta_E+1}\right) \\
& + \log_2\left(1+\frac{(1-\alpha_1-\alpha_2)\gamma 3}{(\alpha_1+\alpha_2)\gamma_3+1}\right) \\
& - \sqrt{\frac{1-\left(1+\frac{\alpha_1\gamma_1}{(1-\alpha_1)\beta_E+1}\right)^{-2}}{n}}Q^{-1}(\epsilon) \\
& - \sqrt{\frac{1-\left(1+\frac{(1-\alpha_1)\gamma 2}{\alpha_1\gamma_2+(1-\alpha_1-\alpha_2)\beta_E+1}\right)^{-2}}{n}}Q^{-1}(\epsilon) \\
& - \sqrt{\frac{1-\left(1+\frac{(1-\alpha_1-\alpha_2)\gamma 3}{(\alpha_1+\alpha_2)\gamma_3+1}\right)^{-2}}{n}}Q^{-1}(\epsilon)
\end{aligned}
$$

$$
\begin{aligned}
\text{s.t.} \quad & \alpha_1 \leq 1 \\
& \alpha_1 \geq 0 \\
& \alpha_2 \leq 1 \\
& \alpha_2 \geq 0 \\
& \alpha_1+\alpha_2 \leq 1 \\
& \alpha_1(-\gamma_1-\beta_E S^t+\beta_E) \leq (-S^t+1)\beta_E \\
& \alpha_1(\gamma_2-\beta_E)S^t+\alpha_2(-\gamma_2-\beta_E S^t) \leq -S^t(\beta_E+1) \\
& \alpha_1\gamma_3(1+S^t)+\alpha_2\gamma_3(1+S^t) \leq -S^t+\gamma_3
\end{aligned}
\tag{8}
$$

4 Results

In this section, we compare the effectiveness of techniques A and B for satisfying quality of service (QoS) restrictions. Following that, we compare the achievable rates with fixed and optimized power allocation coefficients.

MATLAB was employed to execute the optimization procedure, through the fmincon function. Parameter values are: blocklength $n = 400$ bits, error probability $\epsilon = 10^{-6}$, transmit SNR $\rho = 30$ dB, channel estimation error is $\sigma_E^2 = -15$ dB, except otherwise stated.

To facilitate a comparison of the results obtained by approaches A and B, a conversion between target SNR and Rate was necessary. To accomplish this task, we employed equations 4 and 5.

Fig. 1 illustrates the optimal power allocation coefficients required to achieve the maximum sum rate, as determined by approaches A and B. In this scenario, we varied the target rate (or SNR), which represents the minimum necessary for maintaining a certain QoS. It is worth recalling that the target minimum rate is equal for all users. Each user is represented by a different color, the solid line represents the results obtained from approach A and the dashed line from B. The results reveal that as the target rate becomes more stringent, a greater allocation of power is directed towards the weakest user. Approach B notably assigns more power to user 3 and less to 1. The explanation for this discrepancy can be analyzed in the subsequent figure.

Fig. 2 displays the users' rate and sum rate obtained with the optimal power allocation from Fig. 1. Once more, the solid line corresponds to the results derived from approach A, while the dashed line represents those from approach B. Additionally, the rate of each user and the sum rate are portrayed using different colors. As the target rate condition becomes more stringent, the sum rate

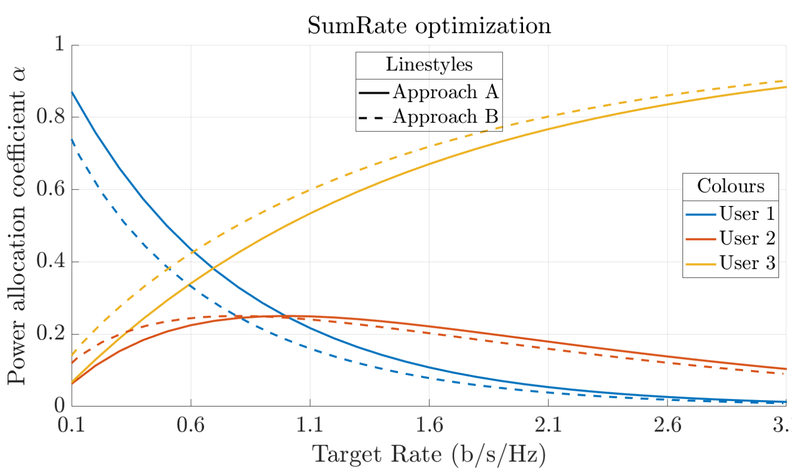

Fig. 1. Optimum power allocation coefficients obtained through approaches A and B

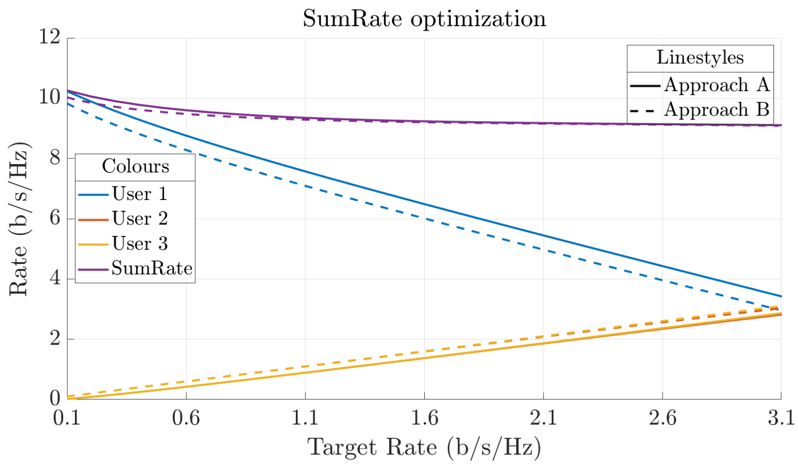

Fig. 2. Sum rate and users' rate calculated with the optimum power allocation coefficients through approaches A and B.

experiences a decrease of approximately 12%. It is observed that the sum rate achieved through the algorithm in approach A appears to exhibit slightly superior performance. This outcome aligns with our prior analysis, indicating that this improvement comes at the expense of allocating less power to the weakest user.

As previously mentioned, approach A produces an inflated rate by overlooking the constraints imposed by the finite blocklength regime in the optimization process. This raises the question of whether the obtained rate for the weaker users meets the true target rate. Fig. 3 provides an alternative visualization of the per-user rates for the sake of a more suitable visualization. The dash-dotted line is the minimum target rate that satisfies QoS and the solid lines are the rates for users 2 and 3 achieved through the algorithm in approach A. The results show clearly that the weaker users do not fulfill the true minimum rate requirements.

To estimate the impact of optimization with respect to a baseline fixed power allocation scheme, we computed the individual users' rates and the overall sum rate with both fixed and optimized power allocation coefficients across various transmit signal-to-noise ratios (ρ). We employed the algorithm outlined in approach B as it was proven to be more reliable. The obtained power allocation coefficients change dynamically, as they represents the optimal values for each transmit SNR. To ensure that we do not set a more restrictive condition for optimization, the target SNR was specified as the minimum SNR obtained from fixed α.

These results are displayed in Figure 4, where the fixed power allocation coefficients for users 1 through 3 are 0.2, 0.3, and 0.5, respectively. The solid line corresponds to the results derived from fixed α, while the dashed line represents

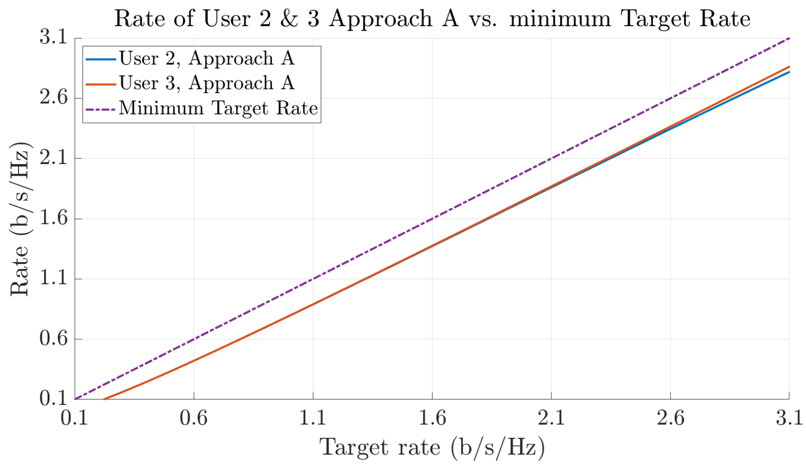

Fig. 3. Minimum target rate and weaker users rate calculated with the optimum power allocation coefficients through approach A.

those from the optimized power allocation coefficient. Additionally, the rate of each user and the sum rate are depicted using different colors.

The optimized α results in a higher total rate than the fixed allocation, although at the expense of providing less power to the weakest user. However, it is worth noting that in this case, rates of users 2 and 3 still satisfy the minimum target rate requirement.

Lastly, in Fig 5, the extension of the results to Effective Capacity are presented. In this instance the delay exponent was set to $\lambda = 0.001$. The optimization goal was to maximize the sum rate, however, this is not necessarily equivalent to maximizing the Effective Capacity (EC), as implied in equation 6. Considering this, it is remarkable that employing the optimized power allocation coefficient results in a higher total effective capacity than the fixed allocation given a relaxed delay requirement.

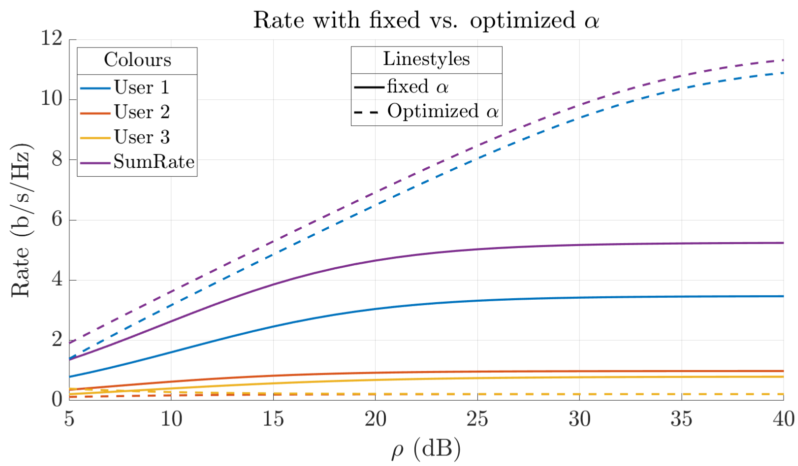

Fig. 4. Sum rate and users' rate calculated with fixed and optimum power allocation coefficients.

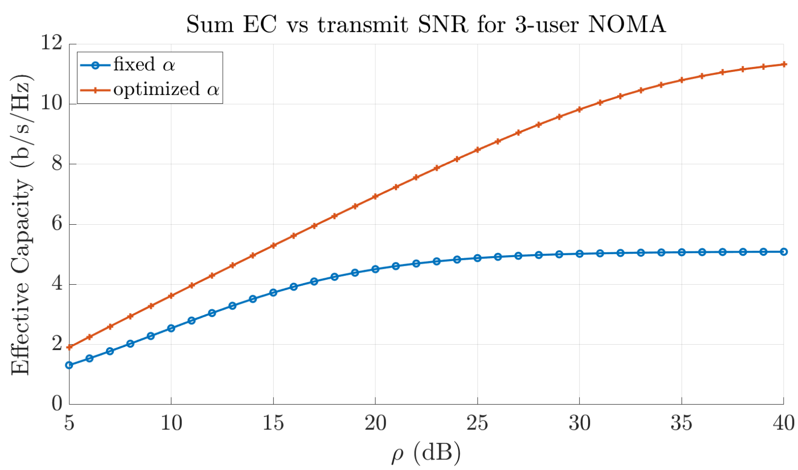

Fig. 5. Sum EC calculated with fixed and optimum power allocation coefficients.

5 Conclusions

This study explored power allocation optimization for low-latency applications in multi-user downlink NOMA systems with imperfect SIC and FBL coding. Two optimization strategies were presented: one based on a modified rate equation and the other one on an achievable signal-to-noise ratio (SNR). The goal was to optimize the effective sum rate while maintaining balanced Quality of Service (QoS) requirements across users.

A comparison between the two optimization approaches encountered discrepancies in power distribution, specifically regarding the weakest user. Approach A tended to allocate less power to the weakest user, sometimes failing to achieve the minimum target rate requirement for QoS. Approach B, which considered the limitations of the finite blocklength regime, demonstrated more reliable performance in fulfilling QoS conditions.

Additional analysis proved that optimized power allocation coefficients resulted in higher total rates than fixed allocation, although possibly delivering less power to the weakest user. Nonetheless, the rates achieved by all users met the minimal goal rate requirement, granting a balanced QoS.

In future work, it is planned to extend the solution to accommodate a greater number of users and to optimize the total Effective Capacity.

References

[1] Y. Yu, H. Chen, Y. Li, Z. Ding, and B. Vucetic, "On the Performance of Non-Orthogonal Multiple Access in Short-Packet Communications," *IEEE Communications Letters*, vol. 22, no. 3, pp. 590–593, 2018. DOI: 10.1109/LCOMM.2017.2786252.

[2] P. Popovski, K. F. Trillingsgaard, O. Simeone, and G. Durisi, "5G Wireless Network Slicing for eMBB, URLLC, and mMTC: A Communication-Theoretic View," *IEEE Access*, vol. 6, pp. 55 765–55 779, 2018. DOI: 10.1109/ACCESS.2018.2872781.

[3] W.-E. Chen, X.-Y. Fan, and L.-X. Chen, "A CNN-based Packet Classification of eMBB, mMTC and URLLC Applications for 5G," in *2019 International Conference on Intelligent Computing and its Emerging Applications (ICEA)*, 2019, pp. 140–145. DOI: 10.1109/ICEA.2019.8858305.

[4] M. Vaezi, A. Azari, S. R. Khosravirad, *et al.*, "Cellular, wide-area, and non-terrestrial iot: A survey on 5g advances and the road toward 6g," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 1117–1174, 2022. DOI: 10.1109/COMST.2022.3151028.

[5] O. L. A. López, H. Alves, R. D. Souza, and M. Latva-Aho, "Finite Blocklength Error Probability Distribution for Designing Ultra Reliable Low Latency Systems," *IEEE Access*, vol. 8, pp. 107 353–107 363, 2020. DOI: 10.1109/ACCESS.2020.3001135.

[6] M. Shehab, E. Dosti, H. Alves, and M. Latva-aho, "On the effective capacity of MTC networks in the finite blocklength regime," in *2017 European Conference on Networks and Communications (EuCNC)*, 2017, pp. 1–5. DOI: 10.1109/EuCNC.2017.7980731.

[7] M. Amjad, L. Musavian, and M. H. Rehmani, "Effective Capacity in Wireless Networks: A Comprehensive Survey," *IEEE Communications Surveys Tutorials*, vol. 21, no. 4, pp. 3007–3038, 2019. DOI: 10.1109/COMST.2019.2929001.

[8] Y. Polyanskiy, "A perspective on massive random-access," in *2017 IEEE International Symposium on Information Theory (ISIT)*, 2017, pp. 2523–2527. DOI: 10.1109/ISIT.2017.8006984.

[9] Y. Wu, X. Gao, S. Zhou, W. Yang, Y. Polyanskiy, and G. Caire, "Massive Access for Future Wireless Communication Systems," *IEEE Wireless Communications*, vol. 27, no. 4, pp. 148–156, 2020. DOI: 10.1109/MWC.001.1900494.

[10] K. Higuchi and A. Benjebbour, "Non-orthogonal Multiple Access (NOMA) with Successive Interference Cancellation for Future Radio Access," *IEICE Transactions on Communications*, vol. E98.B, pp. 403–414, Mar. 2015. DOI: 10.1587/transcom.E98.B.403.

[11] M. C. Fernández Montefiore, G. González, F. J. Lopez-Martinez, and F. Gregorio, "Effective capacity of finite-blocklength NOMA and OMA with heterogeneous latency requirements," in *2021 Argentine Conference on Electronics (CAE)*, 2021, pp. 78–81. DOI: 10.1109/CAE51562.2021.9397566.

[12] M. Amjad, L. Musavian, and S. Aissa, "NOMA versus OMA in Finite Blocklength Regime: Link-Layer Rate Performance," *IEEE Transactions on Vehicular Technology*, pp. 1–1, 2020. DOI: 10.1109/TVT.2020.3037488.

[13] Z. Yang, Z. Ding, P. Fan, and G. K. Karagiannidis, "On the Performance of Non-orthogonal Multiple Access Systems With Partial Channel Information," *IEEE Transactions on Communications*, vol. 64, no. 2, pp. 654–667, 2016. DOI: 10.1109/TCOMM.2015.2511078.

[14] Z. Xiang, X. Tong, and Y. Cai, "Secure transmission for NOMA systems with imperfect SIC," *China Communications*, vol. 17, no. 11, pp. 67–78, 2020. DOI: 10.23919/JCC.2020.11.006.

[15] I. Abu Mahady, E. Bedeer, S. Ikki, and H. Yanikomeroglu, "Sum-Rate Maximization of NOMA Systems Under Imperfect Successive Interference Cancellation," *IEEE Communications Letters*, vol. 23, no. 3, pp. 474–477, 2019. DOI: 10.1109/LCOMM.2019.2893195.

[16] M. Ahmad, S. Baig, H. M. Asif, and K. Raahemifar, "Mitigation of Imperfect Successive Interference Cancellation and Wavelet-Based Nonorthogonal Multiple Access in the 5G Multiuser Downlink Network," *Wireless Communications and Mobile Computing*, vol. 2021, 2021. DOI: https://doi.org/10.1155/2021/8876026.

[17] M. C. F. Montefiore, G. González, F. J. Lopez-Martinez, and F. Gregorio, "Effective rate of multiuser noma systems with imperfect sic for beyond 5g applications," in *2021 XIX Workshop on Information Processing and Control (RPIC)*, 2021, pp. 1–5. DOI: 10.1109/RPIC53795.2021.9648512.