

GA-Optimized BD-RIS-Assisted Network Slicing for Uplink eMBB and URLLC Coexistence

Diana Laura Fernández Duarte, Victoria Dala Pegorara Souto and Richard Demo Souza

Abstract—This paper investigates the impact of Beyond-Diagonal Reconfigurable Intelligent Surfaces (BD-RIS), considering single-, group-, and fully-connected architectures, on the achievable rate region of an uplink system with coexisting eMBB and URLLC services under OMA, NOMA, and RSMA-based RAN slicing. A Genetic Algorithm (GA)-based optimization method is proposed to design the BD-RIS beamforming. Numerical results demonstrate that BD-RIS significantly improves system performance, even under random phase configurations. In particular, the fully-connected architecture achieves the highest eMBB sum rates, while RSMA provides superior interference management capabilities in multiuser scenarios.

Keywords—BD-RIS, network slicing, eMBB, URLLC, RSMA, Genetic Algorithm

I. INTRODUCTION

Sixth-generation (6G) mobile networks are being designed to support the coexistence of heterogeneous services with highly diverse Quality-of-Service (QoS) requirements, which creates major challenges for efficient resource allocation in the Radio Access Network (RAN) [1]. In particular, Enhanced Mobile Broadband (eMBB) targets high data rates, while Ultra-Reliable Low-Latency Communications (URLLC) are intended for critical applications requiring ultra-low latency and very high reliability. In this context, network slicing has emerged as a key solution by partitioning the physical network into multiple logical slices, each tailored to the requirements of a specific service. Traditionally, resource allocation has relied on Orthogonal Multiple Access (OMA), which limits spectral efficiency due to orthogonal resource allocation [2]. As an alternative, Non-Orthogonal Multiple Access (NOMA) enables simultaneous resource sharing, improving spectral efficiency and system capacity. In particular, Power-Domain NOMA (PD-NOMA) multiplexes users with different power levels

over the same resource block, which are subsequently separated at the receiver via Successive Interference Cancellation (SIC) [3].

Several works have investigated the performance of OMA and NOMA for physical-layer RAN slicing. In [4], NOMA is shown to reduce power consumption compared to OMA in a downlink eMBB–URLLC scenario. In [5], MIMO-NOMA is shown to outperform MIMO-OMA under proper power allocation. In [2], NOMA with SIC is shown to be advantageous over OMA in uplink eMBB–URLLC coexistence. In [3], non-orthogonal slicing outperforms orthogonal slicing when URLLC users experience better channel conditions than eMBB users.

Despite these advantages, the performance of NOMA strongly depends on channel asymmetry, which has motivated the study of Rate-Splitting Multiple Access (RSMA) as a more flexible alternative. Unlike NOMA, RSMA improves interference management by splitting messages into common and private components, allowing partial interference decoding while treating the remaining interference as noise [6]. In this context, several works have explored the benefits of RSMA in network slicing scenarios. In [6], RSMA achieves higher data rates in eMBB–URLLC coexistence, even under strict reliability requirements. In [1], RSMA is shown to achieve a larger achievable rate region than NOMA in both eMBB–URLLC and eMBB–mMTC scenarios. In [7], an adaptive scheme among NOMA, RSMA, and puncturing shows that RSMA generally provides higher eMBB rates. In [8], RSMA improves both eMBB rates and URLLC reliability while enhancing user fairness. In [9], an RSMA-based design maximizes the total eMBB rate through optimized power allocation and decoding order, outperforming NOMA. In [10], a hybrid OMA–RSMA scheme is studied under imperfect CSI and imperfect SIC.

However, despite the advantages of network slicing, the coexistence of heterogeneous services still depends on the variability of the wireless channel, which is affected by random fading [11]. In this context, Reconfigurable Intelligent Surfaces (RIS) have emerged as a promising technology by enabling programmable control of the propagation environment. These surfaces, composed of passive reflecting elements, are classified into Diagonal RIS (D-RIS) and Beyond Diagonal RIS (BD-RIS). While D-RIS operate without coupling between elements, BD-RIS enable interconnection through reconfigurable impedance networks, allowing for a more precise control of the wavefront and improving system performance [12].

Recent studies have explored D-RIS-assisted network slicing for heterogeneous service coexistence. In [11], D-RIS is shown to enhance both eMBB rates and URLLC reliability

D. L. F. Duarte e V. D. P. Souto, National Institute of Telecommunications, Santa Rita do Sapucaí - MG, diana.fernandez@dtel.inatel.br, victoria.souto@inatel.br; R. D. Souza, Federal University of Santa Catarina, Florianópolis - SC, richard.demo@ufsc.br. The following research projects have funded this work: Brasil 6G Project with support from RNP/MCTI (01245.020548/2021-07), xGMobile Projects code XGM-AFCCT-2024-4-1-1 and XGM-AFCCT-2026-4-1-1 with resources from EMBRAPII/MCTI (Grant 052/2023 PPI IoT/Manufatura 4.0 and FAPEMIG Grant PPE-00124-23), Bolsas Especiais de IC e Tecnológica nas áreas de Ciência, Tecnologia e Engenharia with resources from FAPEMIG (Grant BIS-00529-25), Ciência por Elas with resources from FAPEMIG (Grant APQ-04523-23), Fomento à Internacionalização das ICTMGs with resources from FAPEMIG (Grant APQ-05305-23), Programa de Apoio a Instalações Multiusuários with resources from FAPEMIG (Grant APQ-01558-24), Redes Estruturantes, de Pesquisa Científica ou de Desenvolvimento Tecnológico with resources from FAPEMIG (Grant RED-00194-23), by CNPq (305021/2021-4, 443974/2024-1, INCT STREAM 409179/2024-8), and by Centro de Competência em Redes 5G e 6G Ref. Finep nº 1060/2 contract nº 0-1-25-0883-00. This work has also been supported by master and doctoral scholarships from CAPES, CNPq, FAPEMIG, xGMobile/EMBRAPII/MCTI.

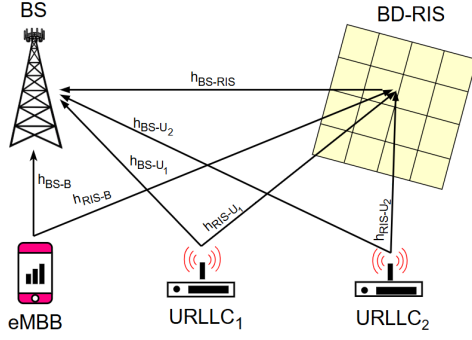


Fig. 1: System Model.

under OMA- and NOMA-based schemes. In [13], a resource allocation algorithm is proposed for D-RIS-assisted NOMA with puncturing, while [14] evaluates different decoding strategies and shows that D-RIS improves the performance of all considered decoders. However, these works are limited to D-RIS with diagonal scattering matrices. To address this limitation, BD-RIS-assisted network slicing was investigated in [15], where single-, group-, and fully-connected architectures were assessed under OMA and NOMA, demonstrating significant performance gains. Nevertheless, the joint integration of RSMA and BD-RIS in this context remains largely unexplored.

Motivated by these gaps, this work investigates an uplink BD-RIS-assisted system with eMBB and URLLC coexistence under OMA, NOMA, and RSMA schemes. The main contributions are summarized as follows: (i) a low-complexity genetic algorithm (GA)-based method is proposed for BD-RIS phase optimization; (ii) it is shown that BD-RIS enlarges the achievable rate region, with the fully-connected architecture maximizing eMBB sum rates and the group-connected architecture providing a favorable performance-complexity trade-off; and (iii) RSMA is shown to achieve the largest achievable rate region when URLLC users experience better channel conditions than eMBB users, whereas OMA becomes preferable in the opposite case.

II. SYSTEM MODEL

We consider an uplink communication system comprising a base station (BS), one eMBB user, and two URLLC users, all equipped with single antennas. Communication is assisted by a BD-RIS with M elements, as illustrated in Fig. 1. The time-frequency resources are arranged in a grid with F frequency resources, indexed by $f \in \{1, \dots, F\}$, and each time slot is divided into S mini-slots, indexed by $s \in \{1, \dots, S\}$. Orthogonal slicing assigns exclusive resources to each service type, while allowing resource sharing among URLLC users. Conversely, non-orthogonal slicing based on NOMA and RSMA enables eMBB and URLLC users to simultaneously share the same radio resources. Owing to stringent latency constraints, URLLC users operate without CSI, whereas perfect CSI is assumed at both the eMBB user and the BS. The eMBB user occupies the entire time slot, while URLLC transmissions are limited to a single mini-slot and spread over multiple frequency resources, with $F_U \leq F$ [1].

The received signal at the BS corresponding to mini-slot s and frequency resource f is given by

$$y_{s,f} = h_{B,f}x_{B,s,f} + \sum_{u \in \{1,2\}} h_{U_u,f}x_{U_u,s,f} + n_{s,f}, \quad (1)$$

where $x_{B,s,f}$ denotes the transmitted signal from the eMBB user, $x_{U_u,s,f}$ represents the signal transmitted by the u -th URLLC user, and $n_{s,f} \sim \mathcal{CN}(0, \sigma^2)$ denotes the additive white Gaussian noise (AWGN) at the BS.

The effective channels between the BS and the eMBB user, and between the BS and the u -th URLLC user, are denoted by $h_{B,f}$ and $h_{U_u,f}$, respectively, and are expressed as

$$h_{B,f} = \mathbf{h}_{BS-RIS,f} \mathbf{\Phi}_B \mathbf{h}_{RIS-B,f} + h_{BS-B,f}, \quad (2)$$

$$h_{U_u,f} = \mathbf{h}_{BS-RIS,f} \mathbf{\Phi}_{U_u} \mathbf{h}_{RIS-U_u,f} + h_{BS-U_u,f}. \quad (3)$$

where $\mathbf{h}_{BS-RIS,f} \in \mathbb{C}^{1 \times M}$ denotes the channel between the BS and the BD-RIS, $\mathbf{h}_{RIS-k,f} \in \mathbb{C}^{M \times 1}$, with $k \in \{B, U_u\}$, represents the channel between the BD-RIS and user k , and $h_{BS-k,f} \in \mathbb{C}^{1 \times 1}$ corresponds to the link between the BS and user k . The scattering matrix of the BD-RIS is denoted by $\mathbf{\Phi}_k$, whose mathematical modeling, considering the single-, group-, and fully-connected architectures, will be detailed later.

All considered links follow independent Rayleigh fading, i.e., $\mathbf{h}_{BS-RIS,f} \sim \mathcal{CN}(0, \beta_{BS-RIS} \mathbf{I})$, $\mathbf{h}_{RIS-k,f} \sim \mathcal{CN}(0, \beta_{RIS-k} \mathbf{I})$, and $h_{BS-k,f} \sim \mathcal{CN}(0, \beta_{BS-k})$. The large-scale path-loss coefficients are given by $\beta_j = C_0 \left(\frac{d_j}{d_0}\right)^{-\alpha_j}$, where j represents any of the aforementioned links, C_0 is the path-loss at the reference distance d_0 , while d_j and α_j denote the distance and the path-loss exponent of link j , respectively.

A. eMBB

The eMBB user adjusts its transmit power according to the instantaneous channel gain using a truncated channel inversion strategy, since perfect CSI is available at both the user and the BS. Specifically, the transmit power is defined as [1]

$$P_B(G_{B,f}) = \begin{cases} \frac{G_{B,f}^{\text{tar}}}{G_{B,f}}, & \text{if } G_{B,f} \geq G_{B,f}^{\min}, \\ 0, & \text{otherwise,} \end{cases} \quad (4)$$

where $G_{B,f} = |h_{B,f}|^2$ denotes the instantaneous channel gain.

According to [1], the minimum threshold required to satisfy the reliability constraint ϵ_B is given by

$$G_{B,f}^{\min} = \Gamma_B \ln \left(\frac{1}{1 - \epsilon_B} \right), \quad (5)$$

where Γ_B denotes the average channel gain of the eMBB user.

The target received power $G_{B,f}^{\text{tar}}$ is obtained by imposing the average power constraint $\mathbb{E}[P_B] = 1$, such that [1]

$$G_{B,f}^{\text{tar}} = \frac{\Gamma_B}{\Gamma \left(0, \frac{G_{B,f}^{\min}}{\Gamma_B} \right)}, \quad (6)$$

where $\Gamma(\cdot, \cdot)$ denotes the upper incomplete Gamma function.

Finally, the achievable eMBB rate is given by [1]

$$r_B = \log_2 \left(1 + \frac{G_{B,f}^{\text{tar}}}{\sigma^2} \right), \quad (7)$$

which remains constant across frequency resources due to the adopted power allocation strategy.

B. URLLC

URLLC users are assumed to transmit with fixed unit power due to the lack of prior CSI [1]. Under this assumption, and considering that the BS decodes URLLC user 1 first followed by URLLC user 2, the corresponding achievable rates are expressed as [1]

$$r_{U_1,f} = \log_2 \left(1 + \frac{G_{U_1,f}}{\sigma^2 + G_{U_2,f}} \right), \quad (8)$$

$$r_{U_2,f} = \log_2 \left(1 + \frac{G_{U_2,f}}{\sigma^2} \right), \quad (9)$$

where $G_{U_1,f} = |h_{U_1,f}|^2$ and $G_{U_2,f} = |h_{U_2,f}|^2$ denote the channel gains of URLLC users 1 and 2, respectively, over frequency resource f .

The outage probability of the URLLC users is defined as

$$P_U = \frac{1}{2} \sum_{u=1}^2 \Pr \left(\frac{1}{F_U} \sum_{f=1}^{F_U} r_{U_u,f} < r_U \right), \quad (10)$$

where r_U denotes the target URLLC rate that satisfies the constraint $P_U \leq \epsilon_U$.

The decoding order is determined by the sum of the mutual information across the F_U frequency resources, given by [1]

$$I_{\text{sum}} = \frac{1}{F_U} \sum_{f=1}^{F_U} \log_2 \left(1 + \frac{G_{U_u,f}}{\sigma^2 + G_{U_w,f}} \right), \quad w \neq u. \quad (11)$$

Finally, the sum rate can be computed as $r_{\text{sum},U} = 2r_U$.

C. BD-RIS Architectures

BD-RIS architectures are classified according to the connectivity of their reconfigurable impedance networks as single-connected, fully-connected, and group-connected [12]. In the single-connected case, also known as D-RIS, each element is independently connected to ground, leading to a diagonal scattering matrix given by $\Phi_k = \text{diag}(e^{j\theta_{k,1}}, \dots, e^{j\theta_{k,M}})$, where $\theta_{k,m} \in (0, 2\pi]$. In the fully-connected architecture, all elements are mutually interconnected, yielding a highly flexible scattering matrix satisfying $\Phi_k = \Phi_k^T$ and $\Phi_k \Phi_k^H = \mathbf{I}_M$, at the expense of increased hardware complexity due to the quadratic growth in the number of impedances. The group-connected architecture provides an intermediate solution by partitioning the M elements into $N_G = M/M_G$ groups of size M_G , resulting in a block-diagonal scattering matrix $\Phi_k = \text{blockdiag}(\Phi_{k,1}, \dots, \Phi_{k,N_G})$, where each block satisfies $\Phi_{k,g} = \Phi_{k,g}^T$ and $\Phi_{k,g} \Phi_{k,g}^H = \mathbf{I}_{M_G}$.

III. PERFORMANCE METRICS

The coexistence of eMBB and URLLC services is evaluated through the achievable sum-rate pair $(r_{\text{sum},B}, r_{\text{sum},U})$ under the reliability constraints (ϵ_B, ϵ_U) for the OMA, NOMA, and RSMA slicing schemes.

1) *OMA*: In OMA-based slicing, the F frequency resources are orthogonally partitioned between services, with F_U resources assigned to URLLC users and $F_B = F - F_U$ to the eMBB user. URLLC users may share their allocated resources, as previously described. Thus, the achievable rates of URLLC users 1 and 2 are given by (8) and (9), respectively, the outage probability by (10), and the URLLC sum rate by $r_{\text{sum},U} = 2r_U$. Since the eMBB user experiences no interference from URLLC transmissions, its achievable rate follows the procedure in Section II-A, yielding (7) and $r_{\text{sum},B} = F_B r_B$.

2) *NOMA*: In NOMA-based slicing, all F frequency resources are shared by URLLC and eMBB services, i.e., $F_U = F_B = F$. Owing to the stringent latency constraints of URLLC, URLLC users are decoded first and therefore experience interference from the eMBB transmission. Accordingly, the target received power of the eMBB user must satisfy [1]

$$G_{B,f}^{\text{tar}} \leq \Gamma_B \left[\Gamma \left(0, \frac{G_{B,f}^{\text{min}}}{\Gamma_B} \right) \right]^{-1}. \quad (12)$$

Under this setting, the eMBB achievable rate is given by (7), and its sum rate follows $r_{\text{sum},B} = F_B r_B$ with $F_B = F$. For URLLC users, the decoding order is defined according to the largest I_{sum} . Assuming that user 1 is decoded before user 2, their achievable rates are expressed as

$$r_{U_1,f} = \log_2 \left(1 + \frac{G_{U_1,f}}{G_{U_2,f} + G_{B,f}^{\text{tar}} + \sigma^2} \right), \quad (13)$$

$$r_{U_2,f} = \log_2 \left(1 + \frac{G_{U_2,f}}{G_{B,f}^{\text{tar}} + \sigma^2} \right). \quad (14)$$

The outage probability is obtained from (10), while the URLLC sum rate is given by $r_{\text{sum},U} = 2r_U$, with $F_U = F$.

3) *RSMA*: In RSMA-based slicing, the URLLC user with the largest I_{sum} splits its message into two sub-messages according to a power-splitting factor $\kappa \in [0, 1]$ [1]. Without loss of generality, URLLC user 1 is assumed to perform rate splitting. In this case, the received signal at the BS over mini-slot s and frequency resource f can be expressed as

$$y_{s,f} = h_{B,f} x_{B,s,f} + h_{U_1,f} \underbrace{\sqrt{\kappa} x_{U_{(1,1)},s,f}}_{\text{1st submessage}} + \underbrace{\sqrt{1-\kappa} x_{U_{(1,2)},s,f}}_{\text{2nd submessage}} + h_{U_2,f} x_{U_2,s,f} + n_{s,f}, \quad (15)$$

where $x_{U_{(1,1)},s,f}$ and $x_{U_{(1,2)},s,f}$ represent the signals associated with the first and second sub-messages of URLLC user 1, respectively.

According to [16], the optimal decoding order is $x_{U_{(1,1)},s,f}$, $x_{U_2,s,f}$, and $x_{U_{(1,2)},s,f}$. Hence, the corresponding achievable rates are given by

$$r_{U_{(1,1)},f} = \log_2 \left(1 + \frac{\kappa G_{U_1,f}}{(1-\kappa)G_{U_1,f} + G_{U_2,f} + G_{B,f}^{\text{tar}} + \sigma^2} \right), \quad (16)$$

$$r_{U_2,f} = \log_2 \left(1 + \frac{G_{U_2,f}}{(1-\kappa)G_{U_1,f} + G_{B,f}^{\text{tar}} + \sigma^2} \right), \quad (17)$$

$$r_{U(1,2),f} = \log_2 \left(1 + \frac{(1 - \kappa)G_{U1,f}}{G_{B,f}^{\text{tar}} + \sigma^2} \right). \quad (18)$$

The total achievable rate of URLLC user 1 is then obtained as the sum of the rates of its two sub-messages, namely,

$$r_{U1,f} = r_{U(1,1),f} + r_{U(1,2),f}. \quad (19)$$

The outage probability is obtained from (10), while the URLLC sum rate is given by $r_{\text{sum},U} = 2r_{U1}$. For the eMBB user, the same procedure adopted in the NOMA-based scheme is applied. Thus, the target received power is given by (12), the achievable rate by (7), and the sum rate by $r_{\text{sum},B} = F_B r_B$, with $F_B = F$.

IV. OPTIMIZATION PROBLEM AND GA-BASED APPROACH

To optimize the BD-RIS beamforming design for the fully-connected and group-connected architectures, we formulate the following optimization problem:

$$\begin{aligned} & \underset{\Phi}{\text{Maximize}} && G_{B,f} \\ & \text{Subject to} && \Phi_{B,g} = \Phi_{B,g}^T, \\ & && \Phi_{B,g} \Phi_{B,g}^H = I_{M_g} \quad \forall g \in \{1, \dots, N_G\}. \end{aligned} \quad (20)$$

This formulation corresponds to the group-connected architecture. When $N_G = 1$, it reduces to the fully-connected case, where all BD-RIS elements are mutually coupled [12].

To solve the optimization problem, we propose a novel GA-based scheme, whose main steps are described as follows.

- 1) Generate an initial population with L random scattering matrices Φ_B .
- 2) Evaluate the fitness of each individual based on $G_{B,f}$.
- 3) Select L parents using tournament selection based on the computed fitness values.
- 4) Generate L offspring from the selected parents. With probability p_{cross} , uniform crossover is performed at the column level; otherwise, the offspring are copied directly from their parents.
- 5) With probability p_m , mutate one element of each offspring matrix Φ_B by replacing it with a randomly generated reflection coefficient.
- 6) Evaluate the fitness of each offspring given by $G_{B,f}$.
- 7) Form the next population by retaining the L_e fittest individuals from the current population and the $L - L_e$ best offspring.
- 8) If the maximum number of iterations N_{it} is reached, return the best scattering matrix Φ_B^* ; otherwise, repeat from Step 2.

V. SIMULATION RESULTS

This section presents numerical results obtained via Monte Carlo simulations in MATLAB® to evaluate the achievable rate region ($r_{\text{sum},B}, r_{\text{sum},U}$) for OMA-, NOMA-, and RSMA-based network slicing assisted by BD-RIS. The system operates over $F = 10$ frequency resources. The BS and BD-RIS are located at (0, 0) m and (102, 2) m, respectively. The path-loss exponents are set as $\alpha_{\text{BS}-U_u} = 3.5$, $\alpha_{\text{RIS}-U_u} = 1.5$, $\alpha_{\text{BS}-B} = 3.5$, $\alpha_{\text{BS}-\text{RIS}} = 1.5$, and $\alpha_{\text{RIS}-B} = 2.5$. The BD-RIS

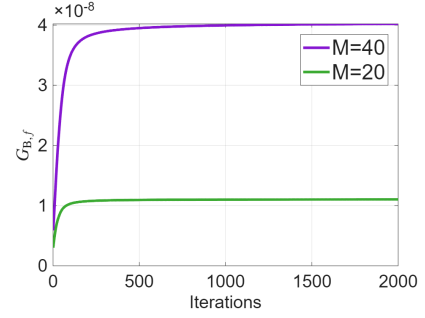


Fig. 2: GA convergence analysis for $M \in \{20, 40\}$

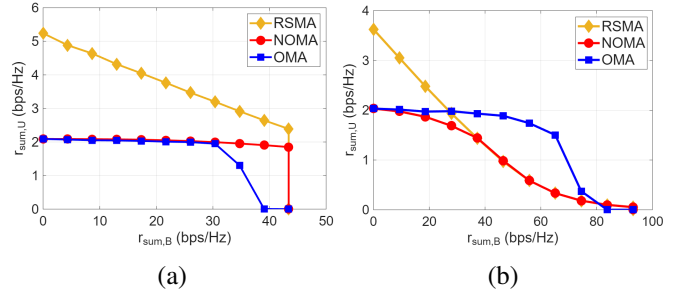


Fig. 3: Achievable rate regions under two deployment scenarios: (a) $(x_b, y_b) = (60, -2)$ m, $(x_u, y_u) = (100, -2)$ m; (b) $(x_b, y_b) = (100, -2)$ m, $(x_u, y_u) = (60, -2)$ m.

has $M = 40$ elements, the noise power is $\sigma^2 = -80$ dBm, and the reliability constraints are $\epsilon_U = 10^{-5}$ and $\epsilon_B = 10^{-3}$. The BD-RIS phase shifts are optimized using the proposed GA with $L = 100$, $N_{\text{it}} = 2000$, $p_{\text{cross}} = 0.9$, $p_m = 0.001$, and $L_e = 5$. Two deployment scenarios are considered: in Scenario 1, the eMBB user is located at $(x_b, y_b) = (60, -2)$ m and both URLLC users at $(x_u, y_u) = (100, -2)$ m; in Scenario 2, these positions are interchanged.

To demonstrate the robustness and convergence of the proposed solution, Fig. 2 depicts the evolution of the instantaneous channel gain $G_{B,f}$ over the GA iterations for different values of M under a fully-connected BD-RIS architecture. The results show that, for all considered values of M , the algorithm converges steadily to high-quality solutions, confirming its effectiveness for BD-RIS beamforming optimization. In addition, larger values of M yield higher channel gains, highlighting the beneficial impact of increasing the number of BD-RIS elements on system performance.

Fig. 3a shows the achievable rate region for Scenario 1 considering a fully-connected BD-RIS ($N_G = 1$). Compared with the case without BD-RIS,¹ the URLLC sum rate increases for all multiple access schemes at an eMBB sum rate of approximately 27 bps/Hz. The largest improvement is achieved by RSMA, with a gain of about 2.85 bps/Hz. This improvement is obtained even with random BD-RIS phase configurations for URLLC users, due to their stringent latency constraints. The BD-RIS also enhances eMBB performance, increasing the maximum sum rate $r_{\text{sum},B}$ from approximately 33.8 bps/Hz to 43.4 bps/Hz. In terms of multiple access

¹The case without BD-RIS was also evaluated, but the corresponding curves are omitted for visual clarity.

TABLE I: $r_{\text{sum},B}$ for different BD-RIS topologies under varying user locations

$(x_b, y_b) / (x_u, y_u)$	$N_G = 1$	$N_G = 2$	$N_G = 5$	$N_G = 40$
(60,-2) / (100,-2)	43.29	43.28	43.12	41.54
(100,-2) / (60,-2)	92.87	92.78	92.24	86.69

strategies, NOMA outperforms OMA at high eMBB sum-rate values because the disparity in user channel conditions favors SIC. RSMA further outperforms both schemes due to its more flexible interference management via rate splitting. These results agree with [1] for the case in which URLLC users experience more favorable channel conditions than the eMBB user.

Fig. 3b presents the results for Scenario 2, in which the user positions are interchanged. In this case, the URLLC sum rate increases only slightly, by about 0.1 bps/Hz, relative to the no-RIS case, since the URLLC users are farther from the BD-RIS. In contrast, the maximum eMBB sum rate increases significantly, reaching approximately 93 bps/Hz with a fully-connected BD-RIS. Under this configuration, OMA outperforms NOMA over most of the achievable rate region because the eMBB user has more favorable channel conditions, causing URLLC users to be decoded under strong interference from the eMBB transmission. At high eMBB sum-rate values, however, NOMA becomes advantageous due to frequency diversity. RSMA outperforms NOMA at low eMBB sum-rate values owing to its greater decoding flexibility, while at high eMBB sum-rate values its performance approaches that of NOMA. Consequently, OMA outperforms RSMA for $28 \leq r_{\text{sum},B} \leq 80$ bps/Hz. These trends are consistent with the results reported in [1], despite the different simulation settings.

To further assess the impact of the BD-RIS topology on system performance, we consider single-connected ($N_G = 40$), group-connected ($N_G \in \{2, 5\}$), and fully-connected ($N_G = 1$) architectures. Table I reports the maximum eMBB sum rate achieved in Scenarios 1 and 2. In both scenarios, the fully-connected architecture outperforms the single-connected one due to its additional degrees of freedom for wavefront control. This gain is more significant in Scenario 2, where the eMBB user is closer to the BD-RIS, yielding an improvement of approximately 6.4 bps/Hz over the single-connected case. Moreover, the group-connected architecture achieves performance close to that of the fully-connected architecture while substantially reducing implementation complexity.

VI. CONCLUSIONS

This paper investigated the uplink coexistence of eMBB and URLLC users under OMA-, NOMA-, and RSMA-based network slicing assisted by BD-RIS. A low-complexity GA-based method was proposed for BD-RIS beamforming design, providing a flexible optimization framework for group- and fully-connected architectures without requiring closed-form solutions. Simulation results showed that BD-RIS enlarges the achievable rate region in all considered scenarios, even with random phase configurations. The fully-connected architecture achieved the highest eMBB sum rates, whereas the group-connected architecture provided comparable performance with lower implementation complexity. Moreover,

RSMA outperformed OMA and NOMA when URLLC users experienced more favorable channel conditions, while OMA became preferable when the eMBB user had better channel conditions. Overall, the results demonstrate that BD-RIS-assisted RAN slicing is a promising solution for efficient uplink resource allocation and eMBB-URLLC coexistence in future wireless networks.

REFERENCES

- [1] Y. Liu, B. Clerckx, and P. Popovski, "Network Slicing for eMBB, URLLC, and mMTC: An Uplink Rate-Splitting Multiple Access Approach," *IEEE Transactions on Wireless Communications*, vol. 23, no. 3, pp. 2140–2152, 2024.
- [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.
- [3] E. N. Tominaga, H. Alves, R. D. Souza, J. Luiz Rebelatto, and M. Latva-aho, "Non-Orthogonal Multiple Access and Network Slicing: Scalable Coexistence of eMBB and URLLC," in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, 2021, pp. 1–6.
- [4] F. Saggese, M. Moretti, and P. Popovski, "NOMA Power Minimization of Downlink Spectrum Slicing for eMBB and URLLC Users," 2021. [Online]. Available: <https://arxiv.org/abs/2106.08847>
- [5] Q. Chen, J. Wang, and H. Jiang, "URLLC and eMBB Coexistence in MIMO Non-orthogonal Multiple Access Systems," 2021. [Online]. Available: <https://arxiv.org/abs/2109.05725>
- [6] E. J. D. Santos, R. D. Souza, and J. L. Rebelatto, "Rate-Splitting Multiple Access for URLLC Uplink in Physical Layer Network Slicing With eMBB," *IEEE Access*, vol. 9, pp. 163 178–163 187, 2021.
- [7] F. Khodakhah, A. Mahmood, P. Österberg, and M. Gidlund, "Adaptive User Pairing With Non-Orthogonal Medium Access Choices for Balanced Coexistence of Mission-Critical and eMBB Services in Cellular IoT," *IEEE Open Journal of the Communications Society*, vol. 6, pp. 5414–5433, 2025.
- [8] M. Katwe, K. Singh, C.-P. Li, S. Prakriya, B. Clerckx, and G. K. Karagiannis, "Enhanced User Fairness and Performance for eMBB-URLLC Uplink Traffic With Rate-Splitting Based Super-Positioning," *IEEE Transactions on Wireless Communications*, vol. 23, no. 9, pp. 12 513–12 530, 2024.
- [9] V. I. Koutsoumpa, N. A. Mitsiou, S. A. Tegos, P. D. Diamantoulakis, T. Lagkas, P. Sarigiannidis, and G. K. Karagiannis, "Network Slicing for eMBB and mMTC with Finite Blocklength Codes," in *2024 13th International Conference on Modern Circuits and Systems Technologies (MOCAST)*, 2024, pp. 1–4.
- [10] A. P. Chrysologou, S. A. Tegos, P. D. Diamantoulakis, N. D. Chatzidi-mantis, P. C. Sofotasios, and G. K. Karagiannis, "On the Coexistence of Heterogeneous Services in 6G Networks: An Imperfection-Aware RSMA Framework," *IEEE Transactions on Communications*, vol. 72, no. 11, pp. 7152–7167, 2024.
- [11] V. D. P. Souto, S. Montejo-Sánchez, J. L. Rebelatto, R. D. Souza, and B. F. Uchôa-Filho, "IRS-Aided Physical Layer Network Slicing for URLLC and eMBB," *IEEE Access*, vol. 9, pp. 163 086–163 098, 2021.
- [12] S. Shen, B. Clerckx, and R. Murch, "Modeling and Architecture Design of Reconfigurable Intelligent Surfaces Using Scattering Parameter Network Analysis," *IEEE Transactions on Wireless Communications*, vol. 21, no. 2, pp. 1229–1243, 2022.
- [13] B. Adhikari, A. Shaharyar Khwaja, M. Jaseemuddin, and A. Anpalagan, "DRL-Leveraged and RIS-Assisted Hybrid Network Slicing for eMBB and URLLC Co-Existence in 6G Systems," *IEEE Open Journal of the Communications Society*, vol. 6, pp. 6156–6176, 2025.
- [14] J. Na, J. Kang, and J. Kang, "Intelligent Reflecting Surface-Assisted Uplink NOMA for eMBB and URLLC Coexistence," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 5, pp. 7406–7411, 2024.
- [15] D. L. F. Duarte, V. Souto, and R. D. Souza, "BD-RIS-Enhanced Physical Layer Network Slicing for URLLC and eMBB Services," in *XLIII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBRT)*, 2025. [Online]. Available: <https://biblioteca.sbirt.org.br/articles/4840>
- [16] Z. Yang, M. Chen, W. Saad, W. Xu, and M. Shikh-Bahaei, "Sum-Rate Maximization of Uplink Rate Splitting Multiple Access (RSMA) Communication," *IEEE Transactions on Mobile Computing*, vol. 21, no. 7, pp. 2596–2609, 2022.