

Rate Region of RIS-Aided URLLC Broadcast Channels: Diagonal versus Beyond Diagonal Globally Passive RIS

Mohammad Soleymani, *Member, IEEE*, Alessio Zappone, *Senior Member, IEEE*, Eduard Jorswieck, *Fellow, IEEE*, Marco Di Renzo *Fellow, IEEE*, and Ignacio Santamaria, *Senior Member, IEEE*,

Abstract—We analyze the finite-block-length rate region of wireless systems aided by reconfigurable intelligent surfaces (RISs), assuming the interference is treated as noise. We consider three nearly passive RIS architectures, including locally passive (LP) diagonal (D), globally passive (GP) D, and GP beyond diagonal (BD) RISs. In a GP RIS, the power constraint is applied globally to the whole surface, while some elements may amplify the incident signal locally. The considered RIS architectures provide substantial performance gains compared with systems operating without RIS. GP BD-RIS outperforms, at the price of increasing the complexity, LP and GP D-RIS as it enlarges the feasible set of allowed solutions. However, the gain provided by BD-RIS decreases with the number of RIS elements. Additionally, deploying RISs provides higher gains as the reliability/latency requirement becomes more stringent.

Index Terms—Broadcast channels, finite block length coding, rate region, reconfigurable intelligent surface.

I. INTRODUCTION

The reconfigurable intelligent surface (RIS) is a promising technology to improve the spectral and energy efficiencies of wireless communication systems [1], [2]. A key feature of RIS

is to operate in a nearly passive manner, without requiring power amplifiers, thus potentially reducing the operational costs, and leading to more environmentally friendly and sustainable communication systems.

A nearly passive RIS can operate in two possible modes [3]: (i) each RIS element operates in a nearly passive mode, hence the amplitude of the reflection coefficient of each RIS element is no greater than one. This architecture is referred to as locally passive (LP); (ii) the total output power of the RIS is no greater than the total input power. In this case, some RIS elements may amplify the incident signal, while the others attenuate it. Globally, however, the whole RIS does not amplify the incident signal. This architecture is referred to as globally passive (GP).

As far as the reconfigurability of the RIS is concerned, the typical architecture assumes that each RIS element is programmed via independent electronic circuits. This architecture is referred to as diagonal (D) RIS. To further increase the number of optimization variables and hence the wave-domain processing capabilities, the RIS elements may be programmed via a fully-connected network of electronic circuits [4]. This architecture is referred to as beyond-diagonal (BD) RIS. A nearly passive BD-RIS is Pareto-optimal from the information-theoretic point of view [5]. The price to pay is, however, a higher implementation complexity and a higher power consumption, compared with D-RIS [6].

A key requirement for sixth-generation (6G) networks is to enhance the link reliability and to reduce the transmission latency [7]. RIS has been shown to be a suitable technology for increasing the reliability and, at the same time, reducing the latency of wireless systems [8]–[12]. For instance, the authors of [8] have shown that a simultaneously transmitting and reflecting (STAR) RIS improves the latency and reliability in multi-cell broadcast channels (BCs). The authors of [10] have proposed an optimization framework for 1-layer rate splitting multiple access in RIS-aided multiple-input single-output (MISO) ultra-reliable and low latency communication (URLLC) systems. The authors of [11] have shown that RIS improves the performance of a two-user single-antenna BC with non-orthogonal multiple access (NOMA). No paper in the literature has, however, considered GP RISs for URLLC applications.

In this paper, we aim at analyzing the achievable rate region of an RIS-aided MISO BC, assuming a finite block length (FBL) coding and treating interference as noise (TIN) as decoding strategy. We devise an algorithm for optimizing the rate region for FBL coding, by establishing a link between

Mohammad Soleymani is with the Signal & System Theory Group, Universität Paderborn, Germany (email: mohammad.soleymani@uni-paderborn.de).

Alessio Zappone is with the Department of Electrical and Information Engineering, University of Cassino and Southern Lazio, 03043 Cassino, Italy, and also with CNIT, 43124 Parma, Italy (e-mail: alessio.zappone@unicas.it)).

Eduard Jorswieck is with the Institute for Communications Technology, Technische Universität Braunschweig, 38106 Braunschweig, Germany (email: jorswieck@ifn.ing.tu-bs.de).

Marco Di Renzo is with Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des Signaux et Systèmes, 91192 Gif-sur-Yvette, France (email: marco.di-renzo@universite-paris-saclay.fr).

Ignacio Santamaria is with the Department of Communications Engineering, Universidad de Cantabria, Spain (email: i.santamaria@unican.es).

The work of A. Zappone was funded by the European Union - Next-GenerationEU under the project NRRP RESTART, RESEARCH and innovation on future Telecommunications systems and networks, to make Italy more smART PE 00000001 (CUP D43C22003080001) [MUR Decree n. 341- 15/03/2022] - Cascade Call launched by SPOKE 3 POLIMI: “SPARKS” project, and project PRIN GARDEN (CUP H53D23000480001). The work of E. Jorswieck was supported by the Federal Ministry of Education and Research (BMBF, Germany) through the Program of Souverän. Digital. Vernetzt. joint Project 6G-RIC, under Grant 16KISK031, and by European Union’s (EU’s) Horizon Europe project 6G-SENSES under Grant 101139282. The work of M. Di Renzo was supported in part by the European Commission through the Horizon Europe project titled COVER under Grant 101086228, the Horizon Europe project titled UNITE under Grant 101129618, and the Horizon Europe project titled INSTINCT under Grant 101139161, as well as by the Agence Nationale de la Recherche (ANR) through the France 2030 project titled ANR-PEPR Networks of the Future under Grant NF-PERSEUS 22-PEFT-004, and by the CHIST-ERA project titled PASSIONATE under Grants CHIST-ERA-22-WAI-04 and ANR-23-CHR4-0003-01. The work of I. Santamaria was funded by MCIN/AEI 10.13039/501100011033, under Grant PID2022-137099NB-C43 (MADRID), and by the Horizon Europe project 6G-SENSES under Grant 101139282.

the rate region and the signal-to-interference-plus-noise ratio (SINR) region, which can be applied if the FBL rate increases monotonically with the SINR. In addition, we compare the average max-min rates of three RIS architectures: LP D-RIS, GP D-RIS, and GP BD-RIS. We show that GP BD-RIS outperforms, in general, the other RIS architectures, at the price of increasing the hardware and computational complexities. In addition, the performance gain provided by BD-RIS is shown to decrease as the number of RIS elements increases.

Notations: The trace of a square matrix $\mathbf{X}$ is denoted as $\text{Tr}(\mathbf{X})$. The complex Gaussian random variable x with mean μ_x and variance σ_x^2 is represented as $\mathcal{CN}(\mu_x, \sigma_x^2)$. The mathematical expectation is denoted as $\mathbb{E}\{\cdot\}$. The real part of a complex variable x is denoted as $\Re\{x\}$. The $M \times M$ identity matrix is denoted as $\mathbf{I}_M$. The Hermitian and transpose of vector/matrix $\mathbf{x}/\mathbf{X}$ are represented by $\mathbf{x}^H/\mathbf{X}^H$ and $\mathbf{x}^T/\mathbf{X}^T$, respectively. The big-O notation is represented by $\mathcal{O}$.

II. SYSTEM MODEL

We consider a single-cell BC with an N -antenna base station (BS), serving K single-antenna users, as shown in Fig. 1. An M -element RIS is considered to aid the transmission between the BS and the users. Global perfect channel state information is assumed. The transmit signal of the BS is $\mathbf{x} = \sum_{k=1}^K \mathbf{w}_k s_k \in \mathbb{C}^{N \times 1}$, where $s_k \sim \mathcal{CN}(0, 1)$ is the message intended for user k , which is one symbol of a codeword of length n , and $\mathbf{w}_k$ is the corresponding beamforming vector. The messages are independent and identically distributed. The power budget of the BS is denoted by p . The transmit power of the BS is $\mathbb{E}\{\mathbf{x}^H \mathbf{x}\} = \sum_k \mathbf{w}_k^H \mathbf{w}_k = \sum_k \text{Tr}(\mathbf{w}_k \mathbf{w}_k^H)$. The notation $\{\mathbf{w}\} = \{\mathbf{w}_k : \forall k\}$ is used to denote the set of all beamforming vectors.

A. RIS Model

The direct link between the users and the BS is assumed to be blocked by obstacles. Thus, the equivalent channel between the BS and user k is $\mathbf{h}_k(\Phi) = \mathbf{f}_k \Phi \mathbf{F} \in \mathbb{C}^{1 \times N}$, where $\mathbf{F}$ is the channel matrix between the BS and the RIS, $\mathbf{f}_k$ is the channel vector between the RIS and user k , and Φ is the matrix of RIS reflection coefficients. The feasible values of Φ depend on the considered RIS architecture. We analyze three case studies.

1) *Locally Passive Diagonal RIS:* In this case, Φ is a diagonal matrix, whose diagonal elements have unit modulus ($|\phi_{mm}| = 1$). Hence, the feasible set is

$$\mathcal{E}_{LP} = \{\phi_{mn} : |\phi_{mm}| = 1, \phi_{mn} = 0, \forall m \neq n\}. \quad (1)$$

2) *Globally Passive Diagonal RIS:* In this case, Φ is still a diagonal matrix, but the power constraint is applied globally to the whole surface. Specifically, Φ is designed such that the output power is not greater than the input power to the RIS, but some RIS elements may amplify the incident signal [3]. Hence, Φ satisfies the following convex constraint:

$$p_{out} - p_{in} = \text{Tr}(\mathbf{F} \mathbb{E}\{\mathbf{x} \mathbf{x}^H\} \mathbf{F}^H (\Phi^H \Phi - \mathbf{I}_M)) \leq 0, \quad (2)$$

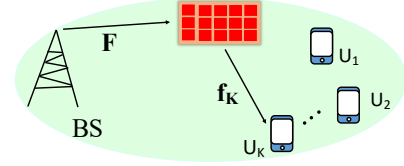


Fig. 1: A MISO BC assisted by an RIS.

where p_{out} and p_{in} are the output and input power to the RIS, respectively. Hence, the feasible set is convex and equal to

$$\mathcal{E}_D = \{\Phi : \Phi = \text{diag}, \text{Tr}(\mathbf{F} \mathbb{E}\{\mathbf{x} \mathbf{x}^H\} \mathbf{F}^H (\Phi^H \Phi - \mathbf{I}_M)) \leq 0\}, \quad (3)$$

with “diag” denoting a diagonal matrix. It is worth noting that the feasible values of Φ depend on the transmission parameters at the BS, i.e., $\{\mathbf{w}\}$, which is not the case for a LP design.

3) *Globally Passive Beyond Diagonal RIS:* In a GP BD-RIS, the matrix Φ can be non-diagonal, but needs to satisfy (2). We assume that Φ is a symmetric matrix, to ease the implementation of BD-RIS [4]. Thus, the feasible set of Φ is

$$\mathcal{E}_{BD} = \{\Phi : \Phi = \Phi^T, \text{Tr}(\mathbf{F} \mathbb{E}\{\mathbf{x} \mathbf{x}^H\} \mathbf{F}^H (\Phi^H \Phi - \mathbf{I}_M)) \leq 0\}. \quad (4)$$

Note that $\mathcal{E}_{LP} \subset \mathcal{E}_D \subset \mathcal{E}_{BD}$, which means that the optimal solution of GP BD-RIS cannot perform worse than any solutions obtained for LP D-RIS and GP D-RIS.

B. Signal Model and Rate Definition

The signal at user k is $y_k = \mathbf{h}_k(\Phi) \sum_{i=1}^K \mathbf{w}_i s_i + n_k$, where $n_k \sim \mathcal{CN}(0, \sigma^2)$ is the additive noise. Employing TIN, the rate of user k using the normal approximation is [13], [14]

$$r_k = \ln(1 + \gamma_k) - Q^{-1}(\epsilon) \sqrt{\frac{v_k}{n}}, \quad (5)$$

where v_k is the channel dispersion of user k , n is the packet length, Q^{-1} is the inverse of the Q-function, ϵ is the maximum tolerable decoding error probability, and γ_k is the signal-to-interference-plus-noise ratio (SINR) at user k given by

$$\gamma_k = \frac{|\mathbf{h}_k(\Phi) \mathbf{w}_k|^2}{\sigma^2 + \sum_{i \neq k} |\mathbf{h}_k(\Phi) \mathbf{w}_i|^2}. \quad (6)$$

An achievable value for the channel dispersion of Gaussian signals in interference-limited systems is $v_k = 2 \frac{\gamma_k}{1 + \gamma_k}$ [14].

Remark 1. In (5), ϵ specifies the reliability requirement, as it denotes the maximum tolerable decoding error probability. Furthermore, the latency is specified by the finite block length n , as a more restrictive latency constraint corresponds to a shorter block length [12], [15].

Lemma 1. The rate function in (5) is increasing in γ_k provided that $\gamma_k \geq \bar{\gamma} = \frac{1}{2}(\sqrt{1 + 2c^2} - 1)$, with $c = Q^{-1}(\epsilon)/\sqrt{n}$.

Proof. The rate function in (5) can be written as $f(\gamma) = \ln(1 + \gamma) - c \sqrt{\frac{2\gamma}{1 + \gamma}}$, where we have suppressed the index k as it is not necessary for this proof. Computing the first-order derivative of f and setting it to be non-negative yields the condition $\sqrt{2\gamma(1 + \gamma)} \geq c$, which, since both γ and c are non-negative quantities, is equivalent to the condition $\gamma \geq \bar{\gamma}$, with $\bar{\gamma}$ being the unique non-negative solution of the equation $2\gamma^2 + 2\gamma - c^2 = 0$. Through simple algebra, it is found $\bar{\gamma} = \frac{1}{2}(\sqrt{1 + 2c^2} - 1)$. $\square$

C. Problem Statement

We aim at deriving the achievable rate region of a MISO BC channel assisted by an RIS. The rate region is the union of all rates achievable by users, i.e., $\mathcal{R} = \bigcup_{\{\mathbf{w} \in \mathcal{W}, \Phi \in \mathcal{E}\}} \{r_k\}_{\forall k}$, where $\mathcal{W}$ is the feasible set of $\{\mathbf{w}\}$.

Employing the rate-profile technique, the achievable rate region is obtained by solving [16]

$$\max_{\{\mathbf{w}\} \in \mathcal{W}, \Phi \in \mathcal{E}, r} r \quad \text{s.t. } r_k \geq \alpha_k r, \forall k, \quad (7)$$

varying the coefficients α_k for every $\alpha_k > 0$ with $\sum_k \alpha_k = 1$.

Additionally, we define the SINR region as the union of all achievable SINRs, which can be obtained by solving

$$\max_{\{\mathbf{w}\} \in \mathcal{W}, \Phi \in \mathcal{E}, \gamma} \gamma \quad \text{s.t. } \gamma_k \geq \lambda_k \gamma, \forall k, \quad (8)$$

varying the coefficients λ_k for every $\lambda_k > 0$ with $\sum_k \lambda_k = 1$.

Solving (8) requires a much lower complexity than solving (7). In the considered scenario, however, and under mild assumptions, the solution obtained by computing the SINR region is equivalent to the solution obtained by computing the rate region, as we prove in the following lemma.

Lemma 2. Assume $\gamma_k \geq \bar{\gamma}$ for all $k = 1, 2, \dots, K$. Then, every point on the boundary of the rate region corresponds to a point on the boundary of the SINR region, and vice versa.

Proof. To prove the lemma, we use a counter example showing that there cannot exist a point on the boundary of the rate region, which is not on the boundary of the SINR region and vice versa. To this end, assume that $(R_1, R_2, \dots, R_K)$ is on the boundary of the rate region, but it is not on the boundary of the SINR region. It means that $(\gamma_1 = f^{-1}(R_1), \gamma_2 = f^{-1}(R_1), \dots, \gamma_K = f^{-1}(R_K))$ is feasible, but it is not on the boundary of the SINR region, where $f^{-1}(x)$ is the inverse of the rate function (denoted as $f(x)$ from Lemma 1. Indeed, Lemma 1 ensures that $f(x)$ is a strictly increasing function whenever $\gamma_k \geq \bar{\gamma}$ holds. It means that there exists at least one SINR $\gamma'_k > \gamma_k$ for $1 \leq k \leq K$ such that $(\gamma_1, \gamma_2, \dots, \gamma'_k, \dots, \gamma_K)$ is feasible, implying that the rates $(R_1, R_2, \dots, f(\gamma'_k) > R_k, \dots, R_K)$ are achievable. Hence, $(R_1, R_2, \dots, R_K)$ cannot be on the boundary of the rate region, which is not correct. Similarly, it can be shown that every point on the boundary of the SINR region is associated with a point on the boundary of the rate region. $\square$

Remark 2. It is important to note that the condition $\gamma_k \geq \bar{\gamma}$ is quite realistic. Indeed, it typically holds for all practical operating conditions, as otherwise the SINR would be too low for ensuring a satisfactory system operation [8, Lemma 2].

Remark 3. The results in Lemma 2 can be extended to any multi-user MISO system using TIN, provided that the rate is a strictly increasing function of the SINR. Thus, the proposed method can be applied not only to the normal approximation, but also to other approximations of the rates, as long as they are strictly increasing functions of the SINR.

III. PROPOSED ALGORITHM

The optimization problem in (8) is not convex. We obtain a suboptimal solution of (8) by combining the majorization

minimization (MM), with the alternating optimization (AO) and the generalized Dinkelbach algorithm (GDA) methods. Specifically, we first solve (8) with respect to $\{\mathbf{w}\}$, while keeping Φ fixed. Then, we solve (8) with respect to Φ , while keeping $\{\mathbf{w}\}$ fixed. In the following, we first provide the solution for updating Φ .

A. Optimization of Φ

The problem in (8) when $\{\mathbf{w}^{(t)}\}$ is kept fixed is

$$\max_{\Phi \in \mathcal{E}, \gamma} \gamma \quad \text{s.t. } \frac{|\mathbf{h}_k(\Phi) \mathbf{w}_k^{(t)}|^2}{\sigma^2 + \sum_{i \neq k} |\mathbf{h}_k(\Phi) \mathbf{w}_i^{(t)}|^2} \geq \max\{\bar{\gamma}, \lambda_k \gamma\}, \forall k, \quad (9)$$

where t is the iteration index. The problem in (9) is a multiple-ratio fractional programming (FP) problem in which the numerator and denominator are quadratic and convex functions of Φ . Also, the set $\mathcal{E}$ is not convex for LP D-RIS.

We first handle the FP problem and then solve (9) for each RIS architecture. To this end, we use the convex-concave procedure (CCP) to approximate the numerator of γ_k with a linear lower bound since it is a convex function. That is

$$|\mathbf{h}_k(\Phi) \mathbf{w}_k^{(t)}|^2 \geq \hat{d}_k(\Phi) \triangleq |\mathbf{h}_k(\Phi^{(t-1)}) \mathbf{w}_k^{(t)}|^2 + 2\Re\{\mathbf{h}_k(\Phi^{(t-1)}) \mathbf{w}_k^{(t)} (\mathbf{h}_k(\Phi) \mathbf{w}_k^{(t)} - \mathbf{h}_k(\Phi^{(t-1)}) \mathbf{w}_k^{(t)})^*\}. \quad (10)$$

Substituting (10) in (9), we obtain

$$\max_{\Phi \in \mathcal{E}, \gamma} \gamma \quad \text{s.t. } \frac{\hat{d}_k(\Phi)}{\sigma^2 + \sum_{i \neq k} |\mathbf{h}_k(\Phi) \mathbf{w}_i^{(t)}|^2} \geq \lambda_k \gamma, \forall k, \quad (11a)$$

$$\frac{\hat{d}_k(\Phi)}{\sigma^2 + \sum_{i \neq k} |\mathbf{h}_k(\Phi) \mathbf{w}_i^{(t)}|^2} \geq \bar{\gamma}, \forall k, \quad (11b)$$

which is a non-convex problem, but it can be optimally solved by using the GDA when $\mathcal{E}$ is a convex set. Indeed, (11b) can be rewritten as the convex constraint

$$\hat{d}_k(\Phi) - \bar{\gamma}(\sigma^2 + \sum_{i \neq k} |\mathbf{h}_k(\Phi) \mathbf{w}_i^{(t)}|^2) \geq 0, \quad \forall k. \quad (12)$$

Thus, we can compute a solution of (11) when $\mathcal{E}$ is convex, by iteratively solving the problem

$$\max_{\Phi \in \mathcal{E}} \min \left\{ \hat{d}_k(\Phi) - \mu^{(t,m)} \lambda_k (\sigma^2 + \sum_{i \neq k} |\mathbf{h}_k(\Phi) \mathbf{w}_i^{(t)}|^2) \right\} \quad \text{s.t. (12),} \quad (13)$$

and updating $\mu^{(t,m)}$ as

$$\mu^{(t,m)} = \min_k \left\{ \frac{\hat{d}_k(\Phi^{(t,m)})}{\lambda_k (\sigma^2 + \sum_{i \neq k} |\mathbf{h}_k(\Phi^{(t,m)}) \mathbf{w}_i^{(t)}|^2)} \right\}, \quad (14)$$

where $\Phi^{(t,m)}$ is the initial point at the m -th iteration of the GDA, which is the solution of (13) at the previous iteration.

1) *GP BD-RIS*: To compute Φ for GP BD-RIS, we iteratively solve the convex problem

$$\max_{\Phi} \min \left\{ \hat{d}_k(\Phi) - \mu^{(t,m)} \lambda_k(\sigma^2 + \sum_{i \neq k} |\mathbf{h}_k(\Phi) \mathbf{w}_i^{(t)}|^2) \right\} \quad (15a)$$

$$\text{s.t. } \sum_k \text{Tr}(\mathbf{F} \mathbf{w}_k^{(t)} \mathbf{w}_k^{(t)H} \mathbf{F}^H (\Phi^H \Phi - \mathbf{I}_M)) \leq 0, \quad (15b)$$

$$(12), \quad \Phi = \Phi^T, \quad (15c)$$

by updating $\mu^{(t,m)}$ according to (14).

2) *GP D-RIS*: To compute Φ for GP D-RIS, we solve the convex problem

$$\max_{\Phi} \min \left\{ \hat{d}_k(\Phi) - \mu^{(t,m)} \lambda_k(\sigma^2 + \sum_{i \neq k} |\mathbf{h}_k(\Phi) \mathbf{w}_i^{(t)}|^2) \right\} \quad (16a)$$

$$\text{s.t. } (12), (15b), \phi_{ij} = 0, \forall i \neq j, \quad (16b)$$

by updating $\mu^{(t,m)}$ according to (14). The algorithm for GP D-RIS is similar to that of GP BD-RIS. The only difference is substituting the constraint in (15c) with the convex constraint $\phi_{ij} = 0$ for all $i \neq j$ to make Φ diagonal. Note that the proposed solutions for GP D-RIS and GP BD-RIS converge to a stationary point of (8) since the proposed algorithms for these architectures fulfill the MM convergence conditions [17].

3) *LP D-RIS*: We employ again the lower bound in (10) and the GDA, which yields

$$\max_{\Phi} \min \left\{ \hat{d}_k(\Phi) - \mu^{(t,m)} \lambda_k(\sigma^2 + \sum_{i \neq k} |\mathbf{h}_k(\Phi) \mathbf{w}_i^{(t)}|^2) \right\} \quad (17a)$$

$$\text{s.t. } (12), |\phi_{ii}| = 1, \phi_{ij} = 0, \forall i \neq j, \quad (17b)$$

with $\mu^{(t,m)}$ given in (14). The problem in (18) is not convex since $|\phi_{ii}| = 1$ is not a convex set. To tackle it, we convexify the constraint $|\phi_{ii}| = 1$ as detailed in [8, Sec. IV.B-1]. The algorithm for LP D-RIS converges since it generates a sequence of non-decreasing values of $\min_k \{\gamma_k\}$ [8].

B. Optimization of $\{\mathbf{w}\}$

To compute $\{\mathbf{w}\}$, we employ the linear lower bound in (10) for the numerator of the SINRs and use the GDA to obtain the global optimal solution of the corresponding surrogate optimization problem. Specifically, we obtain $\{\mathbf{w}\}$ by solving

$$\max_{\mathbf{w}} \min \left\{ \tilde{d}_k(\mathbf{w}) - \mu^{(t,m)} \lambda_k(\sigma^2 + \sum_{i \neq k} |\mathbf{h}_k(\Phi^{(t)}) \mathbf{w}_i^{(t)}|^2) \right\} \quad (18a)$$

$$\text{s.t. } \tilde{d}_k(\mathbf{w}) - \bar{\gamma}(\sigma^2 + \sum_{i \neq k} |\mathbf{h}_k(\Phi) \mathbf{w}_i^{(t)}|^2) \geq 0, \forall k, \quad (18b)$$

$$\sum_k \|\mathbf{w}_k\|^2 \leq p, \quad (18c)$$

where $\mu^{(t,m)}$ is given in (14), and

$$\tilde{d}_k(\mathbf{w}) \triangleq |\mathbf{h}_k(\Phi^{(t-1)}) \mathbf{w}_k^{(t-1)}|^2 + 2\Re \left\{ \mathbf{h}_k(\Phi^{(t-1)}) \mathbf{w}_k^{(t-1)} \left(\mathbf{h}_k(\Phi^{(t-1)}) \mathbf{w}_k - \mathbf{h}_k(\Phi^{(t-1)}) \mathbf{w}_k^{(t-1)} \right)^* \right\}. \quad (19)$$

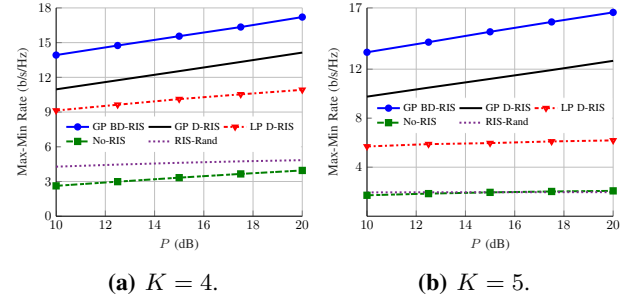


Fig. 2: Max-min rate versus P (Setup: $N = 6$, $n = 256$ bits, $\epsilon = 10^{-5}$, $M = 20$).

C. Computational Complexity

We estimate the computational complexity in terms of the number of complex multiplications to calculate Φ for GP BD-RIS and GP D-RIS. A similar analysis can be conducted for LP D-RIS and $\{\mathbf{w}\}$. To obtain Φ , (16) and (18) need to be solved numerically for GP BD-RIS and GP D-RIS, respectively. To this end, we note that the number of Newton iterations to numerically solve a convex problem is proportional to the square root of the number of constraints [18].

As for $\mathcal{E}_D$ and $\mathcal{E}_{BD}$, two constraints for the feasible set of Φ exist. In addition, there are K constraints for the SINR of each user. Hence, the number of Newton iterations grows with $\sqrt{K}$. To solve (15) or (16) at each Newton step, K channels need to be computed, which approximately requires KM^2N and KMN multiplications for BD-RIS and D-RIS, respectively. To compute the objective function in (15) and (16), the number of multiplications grows with K^2N . Additionally, to compute the constraint (15b), the number of multiplications grows with $KM^2(M+N)$ and KM^2N for GP BD-RIS and GP D-RIS, respectively. Therefore, the computational complexity of one iteration of the GDA for GP BD-RIS and GP D-RIS is of order of $\mathcal{O}(K\sqrt{K}M^2(M+N))$ and $\mathcal{O}(K\sqrt{K}M^2N)$, respectively. In conclusion, the complexity of GP BD-RIS and GP D-RIS increases with M^3 and M^2 , respectively.

IV. NUMERICAL RESULTS

Monte Carlo simulations are utilized to obtain the numerical results. The channel between the BS and the RIS follows a Rician distribution with Rician factor equal to 10 [19, Eqs. (60)-(62)]. The channels between the RIS and the users are in non-line of sight (NLOS) and follow a Rayleigh distribution. The heights of the BS and RIS are equal to 25 meters, while the height of the users is 1.5 meters. The BS is located at the origin, and the distance between the BS and the RIS is 1 meter. The users are randomly located in a square area with side equal to 20 meters, and the distance between the center of the square region and the BS is 130 meters. For comparison, a scenario without the RIS (denoted by **No-RIS**) is considered. In this case, the channels between the BS and the users are NLOS and follow a Rayleigh distribution. The simulation setup is detailed in the captions of each figure. Also, a setup where the reflection coefficients of the RIS are randomly distributed (denoted by **RIS-Rand**) is illustrated as well.

Fig. 2 shows the average value of the max-min rate as a function of the transmit power P . Specifically, we set $\alpha_k = 1/K$ for all k . We see that GP BD-RIS outperforms the other RIS architectures in terms of max-min rate. The price to pay

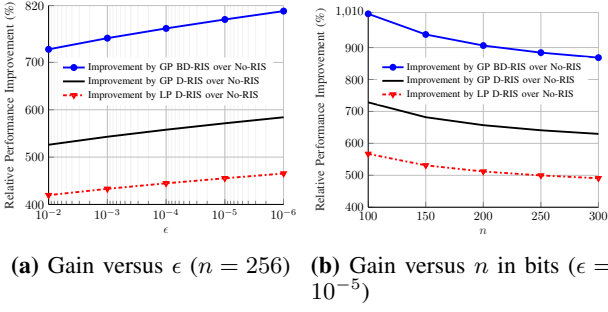


Fig. 3: Gain provided by different RIS architectures (Setup: $N = 3$, $K = 3$, $M = 20$, $P = 10$ dB).

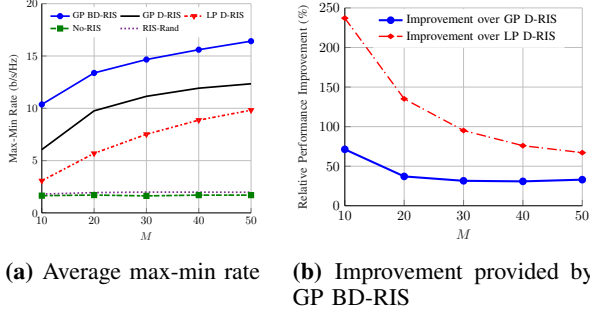


Fig. 4: Impact of M . (Setup: $N = 6$, $K = 5$, $n = 256$ bits, $\epsilon = 10^{-5}$, $P = 10$ dB).

is usually a higher implementation complexity, computational complexity, and reduced energy efficiency [6]. Moreover, when $K = 5$, the gap between the LP and GP architectures (D or BD) is higher. Indeed, in this case, the performance of LP D-RIS is mainly limited by interference, and its max-min rate does not noticeably vary with P . However, GP D-RIS and GP BD-RIS can more effectively manage interference than LP D-RIS when $K = 5$. As expected, we observe that Rand-RIS does not provide any considerable gain, since its elements are not optimized.

Fig. 3 illustrates the gain provided by different RIS architectures, as compared with No-RIS, in terms of average max-min rate, and as a function of ϵ and n . This figure shows that the considered RIS architectures substantially increase the average max-min rate with respect to the No-RIS deployment. Additionally, the gain of each RIS architecture increases when the reliability and latency constraints are more stringent, i.e., when the packets are shorter and when ϵ is lower. In addition, the average max-min rate of GP BD-RIS is approximately 33% and 66% higher than that of GP D-RIS and LP D-RIS, respectively.

Fig. 4 illustrates the average max-min rate as a function of the number M of RIS elements. We see in Fig. 4a that the average max-min rate of RIS-aided systems increases with M only if the RIS elements are optimized according to the proposed schemes. Interestingly, however, the gain provided by GP BD-RIS over GP D-RIS and LP D-RIS decreases as M increases, as depicted in Fig. 4b. Recalling that the number of electronic circuits for D-RIS and fully connected BD-RIS is M and $M(M-1)/2$, respectively, the benefits of BD-RIS needs to account for the implementation complexity and power consumption of the required electronics as well.

V. CONCLUSION

In this paper, we analyzed the rate region of RIS-aided BC channels by considering FBL coding and assuming TIN-based decoding. We proved that the rate region can be obtained from the SINR region, provided that the rate increases monotonically with the SINR. We compared three RIS architectures (LP D-RIS, GP D-RIS, and GP BD-RIS), and have shown that GP BD-RIS, which includes LP D-RIS and GP D-RIS as special cases, provides the highest average max-min rate, at the cost of increasing the implementation and computational complexities. The gain provided by GP BD-RIS reduces, in addition as the number of RIS elements increases.

REFERENCES

- [1] M. Di Renzo *et al.*, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, 2020.
- [2] Q. Wu *et al.*, "Intelligent reflecting surface aided wireless communications: A tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, 2021.
- [3] R. K. Fotock *et al.*, "Energy efficiency optimization in RIS-aided wireless networks: Active versus nearly-passive RIS with global reflection constraints," *IEEE Trans. Commun.*, vol. 72, no. 1, pp. 257–272, 2024.
- [4] H. Li, S. Shen, and B. Clerckx, "Beyond diagonal reconfigurable intelligent surfaces: From transmitting and reflecting modes to single-, group-, and fully-connected architectures," *IEEE Trans. Wireless Commun.*, vol. 22, no. 4, pp. 2311–2324, 2023.
- [5] G. Bartoli *et al.*, "Spatial multiplexing in near field MIMO channels with reconfigurable intelligent surfaces," *IET Signal Process.*, vol. 17, 2023.
- [6] M. Soleymani *et al.*, "Energy efficiency comparison of RIS architectures in MISO broadcast channels," in *IEEE Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*. IEEE, 2024, pp. 701–705.
- [7] C.-X. Wang *et al.*, "On the road to 6G: Visions, requirements, key technologies and testbeds," *IEEE Commun. Surv. Tutor.*, vol. 25, no. 2, pp. 905–974, 2023.
- [8] M. Soleymani, I. Santamaria, and E. Jorswieck, "Spectral and energy efficiency maximization of MISO STAR-RIS-assisted URLLC systems," *IEEE Access*, vol. 11, pp. 70 833–70 852, 2023.
- [9] M. Almekhlafi *et al.*, "Joint resource allocation and phase shift optimization for RIS-aided eMBB/URLLC traffic multiplexing," *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 1304–1319, 2022.
- [10] M. Soleymani *et al.*, "Optimization of rate-splitting multiple access in beyond diagonal RIS-assisted URLLC systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 5, pp. 5063–5078, 2024.
- [11] T.-H. Vu *et al.*, "Intelligent reflecting surface-aided short-packet non-orthogonal multiple access systems," *IEEE Trans. Veh. Technol.*, vol. 71, no. 4, pp. 4500–4505, 2022.
- [12] M. Soleymani *et al.*, "Optimization of the downlink spectral and energy efficiency of RIS-aided multi-user URLLC MIMO systems," *IEEE Trans. Commun.*, 2024, doi: 10.1109/TCOMM.2024.3480988.
- [13] Y. Polyanskiy *et al.*, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.
- [14] J. Scarlett, V. Y. Tan, and G. Durisi, "The dispersion of nearest-neighbor decoding for additive non-Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 63, no. 1, pp. 81–92, 2016.
- [15] Y. Xu, Y. Mao, O. Dizdar, and B. Clerckx, "Max-min fairness of rate-splitting multiple access with finite blocklength communications," *IEEE Trans. Veh. Technol.*, vol. 72, no. 5, pp. 6816–6821, 2023.
- [16] R. Zhang and S. Cui, "Cooperative interference management with MISO beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5450–5458, 2010.
- [17] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, 2017.
- [18] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [19] M. Soleymani, I. Santamaria, and P. J. Schreier, "Improper signaling for multicell MIMO RIS-assisted broadcast channels with I/Q imbalance," *IEEE Trans. Green Commun. Netw.*, vol. 6, no. 2, pp. 723–738, 2022.