

Optimal RIS Partitioning and Power Control for Bidirectional NOMA Networks

Madi Makin¹, Graduate Student Member, IEEE, Sultangali Arzykulov², Member, IEEE,
Abdulkadir Celik³, Senior Member, IEEE, Ahmed M. Eltawil⁴, Senior Member, IEEE,
and Galymzhan Nauryzbayev⁵, Senior Member, IEEE

Abstract—This study delves into the capabilities of reconfigurable intelligent surfaces (RISs) in enhancing bidirectional non-orthogonal multiple access (NOMA) networks. The proposed approach partitions RIS to optimize the channel conditions for NOMA users, improving NOMA gain and eliminating the requirement for uplink (UL) power control. The proposed approach is rigorously evaluated under four practical operational regimes; 1) Quality-of-Service (QoS) sufficient regime, 2) RIS and power efficient regime, 3) max-min fair regime, and 4) maximum throughput regime, each subject to both UL and downlink (DL) QoS constraints. By leveraging decoupled nature of RIS portions and base station (BS) transmit power, closed-form solutions are derived to show how optimal RIS partitioning can meet UL-QoS requirements while optimal BS power control can ensure DL-QoS compliance. Analytical findings are validated by simulations, highlighting the significant benefits that RISs can bring to the NOMA networks in the aforementioned operational scenarios.

Index Terms—Grant-free non-orthogonal multiple access (NOMA), large intelligent surface (LIS), optimization, passive beamforming, reconfigurable intelligent surface (RIS), max-sum rate (MSR).

I. INTRODUCTION

THE advent of reconfigurable intelligent surfaces (RISs) has sparked a paradigm shift in the realm of wireless communication, positioning itself as a revolutionary and highly-promising solution to meet the ever-growing demands of future wireless networks. Utilizing cost-efficient reflecting elements integrated into large planes, RISs possess the capability to dynamically manipulate electromagnetic waves by altering their amplitude and/or phase, thereby enhancing the transmission environment. Through the ability to control

the propagation environment, RISs have been shown to significantly improve the energy efficiency (EE) [1], spectral efficiency (SE) [2], coverage, and overall system throughput [3] of wireless networks, as well as augment the end-to-end performance of multi-hop communications [4]. Furthermore, RISs have been proposed as a supplementary technology for millimeter-wave (mmWave) and THz communications, which are highly susceptible to physical obstacles [5]. In essence, the utilization of RISs promises to revolutionize the way wireless networks operate, offering many benefits that can pave the way for an enhanced communication experience for users.

Non-orthogonal multiple access (NOMA) is a cutting-edge technology that has the potential to revolutionize the field of wireless communications by greatly improving EE and SE while addressing the limitations of its orthogonal counterpart. NOMA achieves this by utilizing the same physical resource blocks (RBs), such as time, frequency, and code, for multiple users. Specifically, power domain (PD)-NOMA utilizes the concept of successive interference cancellation (SIC) to effectively multiplex users with different transmit power and channel gains, thereby significantly enhancing system performance. In addition to its capabilities in traditional wireless networks, NOMA has also been found to have various applications in emerging technologies, such as cognitive radio, unmanned aerial vehicle-assisted cooperative networks, and simultaneous wireless information and power transfer. NOMA is a truly versatile technology with the ability to drive innovation in a wide range of fields, making it a topic of significant interest in the research community.

The integration of RIS and NOMA techniques has the potential to revolutionize the way in which wireless networks operate, as it can further improve key performance indicators of wireless networks. In recent years, there has been a surge in research studies that have sought to explore the capabilities of this combination, with several studies reporting that RIS-enabled NOMA networks can significantly improve SE and EE [6], [7]. Other works have considered dynamic and static RIS configurations for multi-user NOMA schemes [8]. Most of the literature on this topic has focused on centralized scheduling, where a base station (BS) predefines the transmission powers and other network parameters for different users. However, in multi-user PD-NOMA networks, this approach is known to suffer from the issues such as power control and resource allocation complexities, as well as an increase in channel state information (CSI)

Manuscript received 6 March 2023; revised 22 June 2023; accepted 8 August 2023. Date of publication 23 August 2023; date of current version 11 April 2024. This work was supported in part by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan under Grant AP13068122 and in part by the Nazarbayev University under Collaborative Research Program under Grant 11022021CRP1513 (PI: Galymzhan Nauryzbayev). The associate editor coordinating the review of this article and approving it for publication was Y. Liu. (Corresponding author: Madi Makin.)

Madi Makin and Galymzhan Nauryzbayev are with the Department of Electrical and Computer Engineering, School of Engineering and Digital Sciences, Nazarbayev University, Astana 010000, Kazakhstan (e-mail: madi.makin99@gmail.com; galymzhan.nauryzbayev@nu.edu.kz).

Sultangali Arzykulov, Abdulkadir Celik, and Ahmed M. Eltawil are with the Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955, Saudi Arabia (e-mail: sultangali.arzykulov@gmail.com; abdulkadir.celik@kaust.edu.sa; ahmed.eltawil@kaust.edu.sa).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TWC.2023.3306048>.

Digital Object Identifier 10.1109/TWC.2023.3306048

acquisition [9]. These problems in centralized grant-based scheduling approaches must be addressed to reduce complexity at the receiver end of the downlink (DL) and uplink (UL) networks.

To this end, RISs can remedy these problems by creating over-the-air receive power disparity and providing grant-free NOMA (GF-NOMA) [10]. Several recent studies have demonstrated the potential of GF-NOMA schemes, where NOMA users can operate without grant acquisition. The PD-GF-NOMA approach aims to exclude power control and CSI acquisition, especially at the UL transmission, allowing battery and computational power-limited user equipments (UEs) to access available resources without grant acquisition. Accordingly, this paper focuses on a RIS-empowered GF-NOMA method that utilizes the power reception disparity to employ novel algorithms to optimize RIS element assignments and power allocation. By doing so, it aims to address the issues associated with centralized grant-based scheduling approaches and pave the way for a more efficient and streamlined communication experience for users.

A. Related Works

1) *RIS-Empowered NOMA Networks*: Recently, the benefits of RIS-NOMA networks have been applied to various emerging technologies, such as mmWave communication [11], multi-user and multi-cell networks [8], [12], and physical layer security [13], [14], [15]. The authors in [6] explored the utilization of a DL priority-oriented RIS-assisted NOMA network and demonstrated that the proposed model yields an improved ergodic capacity. The integration of RIS technology with NOMA poses several optimization challenges, including rate optimization, optimal power allocation, RIS phase configuration, and UE distribution. To address these challenges, the authors in [16] optimized the achievable rate in the RIS-aided DL transmission scenario under practical system configurations and developed a resource allocation algorithm to maximize the average sum-rate. Additionally, the authors in [17] proposed a RIS-enabled PD-NOMA scheme to achieve practical channel gain differences within NOMA users, and revealed that RIS technology can assist in enlarging channel gain differences, and the effectiveness of the proposed difference-of-convex algorithm in minimizing transmit power.

Furthermore, the authors in [18] investigated a RIS-empowered NOMA-aided backscatter communication system and jointly optimized the power of reflection coefficients and phase shifts to achieve superior performance compared to the conventional NOMA and OMA backscattering systems, through a low complexity algorithm. Additionally, in [7], the authors studied a RIS-deployed NOMA mobile edge computing network, intending to minimize the sum energy consumption through optimization of transmission rate, power control, and passive phase shifters. The results demonstrated the effectiveness of the proposed algorithm in increasing EE of the system. Similarly, in [19], a two-step machine learning (ML) method was proposed to jointly optimize power allocation, phase shift, and user localization, while, in [20], ML-based algorithms were applied to the RIS-empowered

NOMA networks for phase shift design and NOMA user partitioning, and, in [21], a deep reinforcement learning (DRL) approach was applied to optimize the network.

2) *Grant-Free NOMA Networks*: The fundamental principle of NOMA is to simultaneously transmit multiple streams of information over the same radio RBs by utilizing advanced multi-user signal decoding algorithms at the receiver [22]. Various methods have been proposed in the literature to separate users in NOMA networks, such as spreading, scrambling, interleaving, or multiple domains [9]. These methods all employ a common approach of multiplexing users over the same time-frequency RBs with the use of differentiated parameters. However, the existing literature primarily focuses on centralized scheduling plans, where in DL transmission, the BS predetermines the power levels of each user, which leads to a significant increase in signaling overhead. In UL traffic, users must be aware of the channel quality of other users in order to align their transmission power accordingly. In light of these challenges, the use of GF-NOMA has become an indispensable technique for reducing signaling overhead and computational complexity at the receiver [23]. A variety of techniques for GF-NOMA, such as multiple access-, compressed sensing-, compute and forward-based, have been extensively discussed in the literature.

The researchers have extensively explored the potential of GF-NOMA to enhance the performance of wireless communication systems. In [24], the authors investigated the use of ALOHA GF-NOMA schemes to mitigate user complexity and increase throughput, while [10] proposed a collision avoidance method for GF-NOMA. The utilization of GF-NOMA has also been studied from a stochastic geometry perspective for semi-GF and compressed sensing-based schemes in [25] and [26], respectively. Furthermore, the authors in [27] applied the index modulation method in the UL GF-NOMA transmission scenario. Additionally, the application of artificial intelligence (AI) techniques has been explored to optimize the performance of GF-NOMA. In [28], the authors proposed a deep learning approach to solve the variational optimization problem for GF-NOMA, while, in [29], random and structured sparsity learning was utilized to reduce users' signaling overhead. Finally, the use of DRL was proposed in [30] and [31] to optimize the transmit power in semi- and full GF-NOMA schemes, respectively.

The GF-NOMA scheme can also be implemented using the method of RIS partitioning, where a specific number of RIS elements are dedicated to relaying the superimposed signal from BS to each NOMA user [32], [33], [34]. Most of the aforementioned studies focus on either uplink or downlink transmission scenarios for RIS-assisted systems. Moreover, the majority of the works above do not leverage the virtual RIS partitioning approach to eliminate the need for uplink power control required by PD-NOMA schemes.

B. Main Contributions

To the best of our knowledge, this paper is the first to present an optimization approach for RIS partitioning and power control in bidirectional NOMA networks using the

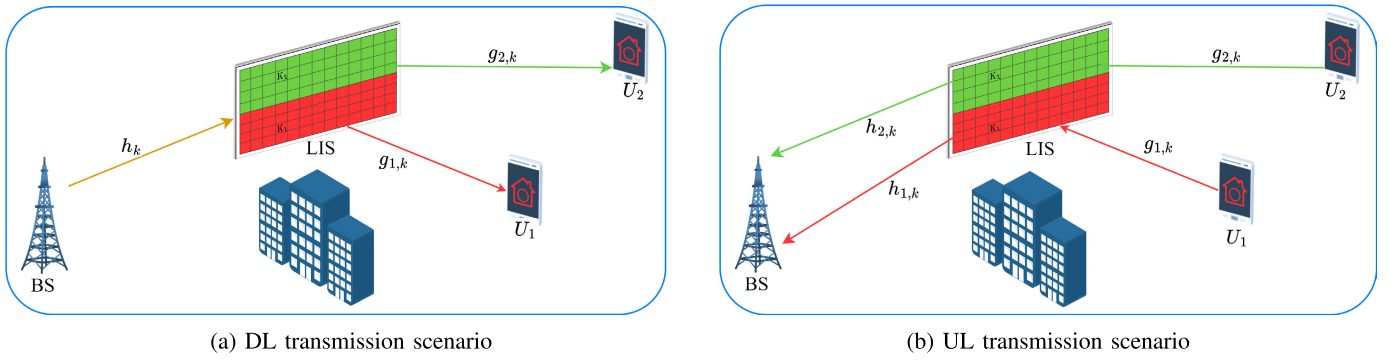


Fig. 1. The system model of RIS-partitioning in DL and UL GF-NOMA networks.

concept of GF-NOMA.¹ Our research aims to analyze and optimize bidirectional NOMA networks through the proposal of a novel RIS partitioning approach. By utilizing this concept, we optimize the channel conditions for NOMA users, enhance the NOMA gain, and eliminate the need for uplink power control. We also introduce the concept of RIS-aided GF-NOMA by designing and employing innovative algorithms for optimal power control and RIS partitioning. The main contributions of this work can be summarized as follows:

- The present study brings to light a pioneering approach that addresses the joint optimization of RIS partitioning and BS power control in order to facilitate bidirectional NOMA traffic while concurrently satisfying both DL and UL Quality-of-Service (QoS) demands. The proposed strategy encompasses the GF-NOMA scheme in UL transmission, as the need for power control by UEs is eliminated by manipulating the portion of RIS assigned to improve channels for each UE.
- The proposed system model is evaluated under four key practical operational traffic regimes: 1) QoS sufficient regime, 2) RIS and power efficient regime, 3) max-min fair regime, and 4) maximum throughput regime, all subject to the UL-QoS and DL-QoS constraints. We derive optimal RIS partitioning and BS power allocation in closed forms (CFs) for the first three objectives. By leveraging the decoupled nature of these variables, we demonstrate how RIS partitioning can meet UL-QoS demands, while BS power control is utilized in cases where UL optimal RIS partitioning violates DL-QoS demands. In the maximum throughput regime, our aim is to maximize the weighted sum-rates of bidirectional traffic. To this end, we leverage metaheuristic approaches to find optimal RIS portions and DL power control levels.
- We conduct extensive simulations to validate our analytical findings through comparisons with well-known optimization solvers. The numerical results demonstrate that the derived CF solutions closely match numerical solutions. Additionally, simulations reveal that the proposed approach can satisfy the bi-directional traffic

demands while adhering to the four main objectives outlined above. Even though this paper considers perfect SIC and CSI to provide more insights into the proposed RIS-assisted bi-directional NOMA approach, we also provide the detrimental impacts of practical imperfections.

C. Paper Organization

The rest of the work is organized as follows. Section II introduces and describes the system model and the corresponding transmission protocol for the given RIS-enabled GF-NOMA network. Section III states the optimization problems, such as the RIS and power efficient regime, max-min fairness, and max-sum rate (MSR), considering their solutions' methodology. In Section IV, the corresponding numerical results with detailed discussions are considered for the performance evaluation. Lastly, we provide concluding remarks in Section V.

II. SYSTEM MODEL

A. System Topology

This study explores the bidirectional transmission of a RIS-assisted NOMA network that comprises single antenna BS/UEs and RIS with K passive reflecting elements, as depicted in Fig. 1. The paper adopts the Nakagami- m fading model [35] and the CSI is assumed to be available by using advanced model-based or data-driven channel estimation techniques as described in [36] and [37]. Hence, proposed solutions can serve as an upperbound to practical cases with SIC, CSI, RIS related imperfections. Moreover, the NOMA users are assumed to have limited mobility and experience slow-fading. With CSI at hand, BS is able to compute the optimal RIS phase shifts and transmit them to the RIS controller via a dedicated wireless [1] or wired [38] feedback channel. In NOMA schemes, the power reception disparity is a crucial factor determining the overall performance gain over its orthogonal counterpart. In order to mitigate this, RIS is partitioned to serve multiple users such that each partition is configured to improve the link quality of the relevant user. As a result, each user receives the reflected signal with coherently and non-coherently aligned phases from its own and other users' partitions, respectively. This approach improves the overall system performance by ensuring each user receives a signal optimally tailored to users' specific needs.

¹It is worth noting the importance of reliability and collision issues in GF-NOMA. However, this work focuses on analyzing and optimizing cooperative NOMA networks in the physical layer, and therefore, we do not explicitly address the reliability and collision issues of GF-NOMA in this paper.

B. Bidirectional Transmission Protocol

In the following, we outline the transmission protocol for DL and UL RIS-assisted NOMA networks with two users. It is worth noting that, for the purpose of simplicity, we consider only two NOMA users in this analysis. However, it is important to acknowledge that serving multiple NOMA nodes can pose significant challenges, as the complexity of the SIC process escalates exponentially with an increased number of users, as previously reported in the literature such as [39]. Furthermore, this issue is further exacerbated in the presence of hardware-related SIC imperfections, as highlighted in [40].

1) *DL Transmission*: In the DL transmission (see Fig. 1a), U_i , $i \in \{1, 2\}$, receives the following signal

$$\tilde{y}_i = \underbrace{\left[\sum_{k=1}^{K_1} g_{i,k} h_k \psi_i e^{j\tilde{\theta}_{i,k}} \right]}_{\text{aligned phase}} + \underbrace{\left[\sum_{k=1}^{K_2} g_{i,k} h_k \psi_i e^{j\tilde{\theta}_{r,k}} \right]}_{\text{not aligned phase}} \times \sqrt{P_b} \left[\sqrt{\beta_1} x_1 + \sqrt{\beta_2} x_2 \right] + n_i, \quad (1)$$

where P_b corresponds to the transmit power at BS; $K_i = \lceil \alpha_i K \rceil$ and $K_r = \lceil \alpha_r K \rceil$, with α_i and α_r denoting the RIS elements allocation coefficients with $i, r \in \{1, 2\}$, $i \neq r$ and $\alpha_i + \alpha_r \leq 1$; K corresponds to the overall number of RIS elements deployed, β_1 and β_2 denote the BS power allocation coefficients, $\beta_1 + \beta_2 \leq 1$; $g_{i,k}$ is the channel between user i and RIS portion's k th element while h_k denotes the channel between the k th element and BS; $\psi_i = d_{i,k}^{-\frac{\tau}{2}} d_{k,b}^{-\frac{\tau}{2}} \rho$, where $d_{i,k}$ and $d_{k,b}$ denote the corresponding distances of UE $_i$ -RIS and RIS-BS links, respectively; τ is the path-loss coefficient, and ρ is the reflecting coefficient; n_i corresponds to the additive white Gaussian noise (AWGN) with the variance of σ_i^2 . Finally, $\tilde{\theta}_{l,k}$ denotes the RIS's phase adjusting coefficients for NOMA user l , $\forall l \in \{i, r\}$, in the DL transmission scenario, which can be expressed as $\tilde{\theta}_{i,k} = -(\arg[g_{i,k}] + \arg[h_k])$. The following expression of the received signal is not tractable for analytical purposes; therefore, we propose an approximate representation which can be written as

$$\tilde{y}_i = \underbrace{\left[\alpha_i \sum_{k=1}^K g_{i,k} h_k \psi_i e^{j\tilde{\theta}_{i,k}} \right]}_{\text{aligned phase}} + \underbrace{\left[\alpha_r \sum_{k=1}^K g_{i,k} h_k \psi_i e^{j\tilde{\theta}_{r,k}} \right]}_{\text{not aligned phase}} \times \sqrt{P_b} \left[\sqrt{\beta_1} x_1 + \sqrt{\beta_2} x_2 \right] + n_i, \quad (2)$$

which will be compared with the exact model in the prospective numerical results section. Without loss of generality and for the sake of explanation clarity, we assume that U_1 has a stronger channel. In this case, U_1 detects the signal of U_2 first, removes U_2 's message from the received signal, and only then decodes its own message. On the other hand, U_2 can decode its message by treating the message of U_1 as interference. Subsequently, the generalized signal-to-interference-to-noise ratio (SINR) for each user can be expressed as

$$\tilde{\gamma}_i(\alpha, \beta) = \frac{\beta_i P_b \left| \alpha_i \sum_{k=1}^K \Psi_i e^{j\tilde{\theta}_{i,k}} + \alpha_r \sum_{k=1}^K \Psi_i e^{j\tilde{\theta}_{r,k}} \right|^2}{\xi \beta_r P_b \left| \alpha_i \sum_{k=1}^K \Psi_i e^{j\tilde{\theta}_{i,k}} + \alpha_r \sum_{k=1}^K \Psi_i e^{j\tilde{\theta}_{r,k}} \right|^2 + \sigma_i^2}, \quad (3)$$

where $\alpha = [\alpha_1, \alpha_2]$, $\beta = [\beta_1, \beta_2]$, $\Psi_i = g_{i,k} h_k \psi_i$ and ξ is a decoding order variable given by

$$\xi = \begin{cases} 1, & \text{if } i = 2, \\ 0, & \text{if } i = 1. \end{cases}, \quad (4)$$

which dictates which user cancels or experiences interference, as explained above.

2) *UL Transmission*: Unlike the DL transmission, we impose no power control at UE side. Alternatively, the power control required for UL NOMA is constituted by manipulating RIS partitions. This inherently paves the way for over-the-air power control as power reception disparity is needed by NOMA managed by RIS without explicit power control at the UE side. In light of these discussions, during the UL transmission (see Fig. 1b), the signals transmitted by both users are superposed at BS as follows

$$\hat{y} = \sqrt{P_1} \left[\alpha_1 \sum_{k=1}^K \Upsilon_1 e^{j\hat{\theta}_{1,k}} + \alpha_2 \sum_{k=1}^K \Upsilon_1 e^{j\hat{\theta}_{2,k}} \right] x_1 + \sqrt{P_2} \left[\alpha_2 \sum_{k=1}^K \Upsilon_2 e^{j\hat{\theta}_{2,k}} + \alpha_1 \sum_{k=1}^K \Upsilon_2 e^{j\hat{\theta}_{1,k}} \right] x_2 + n_b, \quad (5)$$

where $\Upsilon_i = g_{i,k} h_{i,k} \psi_i$, P_1 and P_2 are the transmit power of NOMA users, n_b is AWGN with the variance of σ_b^2 ; $h_{i,k}$ is the channel between the definite part of RIS's k th element, passing UE $_i$'s signal; $\hat{\theta}_{i,k}$ is the RIS's phase shifting parameter for the UL transmission, which can be expressed as $\hat{\theta}_{i,k} = -(\arg[g_{i,k}] + \arg[h_{i,k}])$. From (5), the receiving generalized SINRs can be expressed as

$$\hat{\gamma}_i(\alpha) = \frac{P_i \left| \alpha_i \sum_{k=1}^K \Upsilon_i e^{j\hat{\theta}_{i,k}} + \alpha_r \sum_{k=1}^K \Upsilon_i e^{j\hat{\theta}_{r,k}} \right|^2}{\xi^C P_r \left| \alpha_r \sum_{k=1}^K \Upsilon_r e^{j\hat{\theta}_{r,k}} + \alpha_i \sum_{k=1}^K \Upsilon_r e^{j\hat{\theta}_{i,k}} \right|^2 + \sigma_0^2}, \quad (6)$$

where ξ^C is the complement of DL decoding order ξ , describing the key distinction between the DL and UL transmissions. Unlike the DL SINR expression in (3), the UL SINR expression in (6) is a function of α as there is no power control at the UE side. The UL NOMA protocol is performed precisely opposite to DL NOMA. BS performs SIC in the UL scenario, where preserving the difference between the received superimposed signal is vital. Given identical users' transmit powers, such difference is achieved with their corresponding channel gains. Thus, BS receives more substantial signal power from the user with a stronger channel gain. Then, unlike in DL NOMA, the stronger user's message is decoded by treating other messages as interference, while a weak user's message will be decoded after removing the stronger user's message from the received signal. Without loss of generality, the rest of the paper considers unity system bandwidth for the sake of presentation. Therefore, SE and data rate are used interchangeably throughout the paper.

III. PROBLEM STATEMENT AND SOLUTION METHODOLOGY

This section first presents formal problem statements and then develops the proposed solution methodologies. By taking

the DL and UL QoS constraints into account, we focus on optimal RIS partitioning for four main transmission regimes: QoS sufficient regime, the RIS and power efficient regime, the max-min fair regime, and the maximum throughput regime, which are explained in detail next.

A. QoS Sufficient Regime

As an essential background to the next optimization problems, we first focus on the QoS sufficient regime that ensures all QoS constraints are satisfied for all users for both UL and DL transmissions. Solving the QoS sufficient regime helps us understand the feasible regions of RIS portion allocations facilitating the bidirectional traffic, which will eventually provide a basis for the other regimes considered in subsequent subsections. The feasibility problem can be formulated as

$$\begin{aligned} \mathbf{P}_0 : \min_{\alpha_1, \alpha_2} c \\ \text{C}_0^1: \quad \text{s.t. } \hat{\gamma}_i(\alpha, \beta) \geq \hat{\gamma}_{\text{th}}^i, i \in \{1, 2\}, \\ \text{C}_0^2: \quad \check{\gamma}_i(\alpha) \geq \check{\gamma}_{\text{th}}^i, i \in \{1, 2\}, \\ \text{C}_0^3: \quad \alpha_1 + \alpha_2 \leq 1, \\ \text{C}_0^4: \quad \alpha_1 \in [0, 1], \alpha_2 \in [0, 1], \end{aligned} \quad (7)$$

where c is an arbitrary constant, C_0^1 and C_0^2 are the UL and DL QoS constraints, C_0^3 is the total RIS portion allocation constraint, C_0^4 is the domain of optimization variables. The feasible region of $\alpha_i, i \in \{1, 2\}$, can be obtained by finding the roots of inequality constraints in C_0^1 and C_0^2 . Accordingly, the following lemma provides the feasible region of RIS portions that satisfies the UL and DL QoS constraints.

Lemma 1: For given BS power allocation, the QoS sufficient feasible region of $\alpha_i, i \in \{1, 2\}$, is given by

$$\max_{j \in \{1, 2\}} \left\{ \hat{\alpha}_{i,\ell}^j(\beta), \check{\alpha}_{i,\ell}^j(\beta) \right\} \leq \alpha_i \leq \min_{j \in \{1, 2\}} \left\{ \hat{\alpha}_{i,v}^j(\beta), \check{\alpha}_{i,v}^j(\beta) \right\}, \quad (8)$$

where $\hat{\alpha}_{i,\ell}^j(\beta)/\check{\alpha}_{i,\ell}^j(\beta)$ denotes the lower bound (LB) of α_i that satisfies the UE_j 's UL/DL QoS constraints. Likewise, $\hat{\alpha}_{i,v}^j(\beta)/\check{\alpha}_{i,v}^j(\beta)$ represents the upper bound (UB) of α_i that satisfies the U_i 's UL/DL QoS constraints. The feasibility of ranges provided in (8) is subject to the satisfaction of C_0^3 in (7).

Proof: Please see Appendix A. ■

B. RIS and Power Efficient Regime

The RIS and power efficient regime aims at optimizing both RIS portions $\alpha_i, i \in \{1, 2\}$, and power allocation parameters $\beta_i, i \in \{1, 2\}$, to reach minimal use of the BS transmit power and RIS elements subject to the user-specific DL and UL QoS constraints. In this way, the RIS and power efficient regime will enable the efficient use of RIS and power, sparing the remaining BS transmit power and RIS elements for other purposes. Unlike the high transmission power available at the BS side, the battery-limited UEs are generally restricted by low transmission powers. Since we are interested in manipulating the UL power control through RIS partitioning, determining optimal RIS partitions jointly for both UL and DL traffic may

yield infeasible solutions. Accordingly, we propose to manage the RIS and power usage minimization problem in two stages: first, we obtain the minimal RIS element allocations that satisfy the UL QoS constraints. Next, the BS power consumption is minimized while satisfying the DL QoS constraints based on the UL optimal RIS partitions.

The RIS usage optimization problem for the UL scenario can be formulated as

$$\begin{aligned} \mathbf{P}_1 : \min_{\alpha_1, \alpha_2} \alpha_1 + \alpha_2 \\ \text{C}_1^1: \quad \text{s.t. } \hat{\gamma}_i(\alpha) \geq \hat{\gamma}_{\text{th}}^i, i \in \{1, 2\}, \\ \text{C}_1^2: \quad \alpha_1 + \alpha_2 \leq 1, \\ \text{C}_1^3: \quad \alpha_1 \in [0, 1], \alpha_2 \in [0, 1], \end{aligned} \quad (9)$$

where C_1^1 is the UL QoS constraint, $\hat{\gamma}_{\text{th}}^i \triangleq 2^{\hat{R}_{\text{th}}^i} - 1$ is the UL QoS requirement of $U_i, i \in \{1, 2\}$. C_1^2 limits the total portioning to the actual number of elements, and C_1^3 defines the variables domain. It is worth noting that the minimal RIS element usage will be attained when all QoS constraints are satisfied at equality, i.e., $\hat{\gamma}_i(\alpha) = \hat{\gamma}_{\text{th}}^i, i \in \{1, 2\}$, since providing users with higher rates will require higher RIS element usage and violate the objective function. Due to the complexity of C_1^1 , the CF solutions for this problem are presented with neglecting the impact of not-aligned phases. Nevertheless, the CF solutions for the original configuration were derived with the assistance of numerical solvers and will be presented in the proceeding numerical results section. Accordingly, the following lemma provides CF optimal RIS portions that minimize the RIS usage while satisfying the UL QoS constraints.

Lemma 2: The minimum RIS portions satisfying all UL QoS constraints are given by

$$\hat{\alpha}_1 = \sqrt{\frac{\hat{\gamma}_{\text{th}}^2 \left(\sigma_b^2 + \hat{\alpha}_2^2 B_2^2 P_2 \right)}{P_1 B_1^2}}, \quad \hat{\alpha}_2 = \sqrt{\frac{\hat{\gamma}_{\text{th}}^2 \sigma_b^2}{P_2 B_2^2}}. \quad (10)$$

Proof: Please see the proofs and notations in Appendix B. ■

Having obtained both $\hat{\alpha}_1$ and $\hat{\alpha}_2$, the optimization problem for the DL scenario can be likewise defined as

$$\begin{aligned} \mathbf{P}_2 : \min_{\beta_1, \beta_2} \beta_1 + \beta_2 \\ \text{C}_2^1: \quad \text{s.t. } \check{\gamma}_i(\alpha, \beta) \geq \check{\gamma}_{\text{th}}^i, i \in \{1, 2\}, \\ \text{C}_2^2: \quad \beta_1 + \beta_2 \leq 1, \\ \text{C}_2^3: \quad \beta_1 \in [0, 1], \beta_2 \in [0, 1], \end{aligned} \quad (11)$$

where C_2^1 is the DL QoS constraint, C_2^2 limits the total transmit power, and C_2^3 characterizes the variables domain. Similar to C_1^1 , C_2^1 is also satisfied with equality, i.e., $\check{\gamma}_i(\beta, \hat{\alpha}) = \check{\gamma}_{\text{th}}^i, i \in \{1, 2\}$, at the optimal point since providing a DL rate more than the required QoS threshold will result in consuming more power and violate the optimality.

Lemma 3: Using the minimal RIS portions $\hat{\alpha}$, the optimal power allocations minimizing the total power consumption

while satisfying DL QoS constraints are given by

$$\beta_1^* = \frac{\hat{\gamma}_{th}^2 \sigma_1^2}{\alpha_1^2 A_1^2 P_1}, \quad \beta_2^* = \frac{\hat{\gamma}_{th}^2 (\sigma_2^2 + \beta_1^* \alpha_2^2 A_2^2 P_2)}{P_2 \alpha_2^2 A_2^2}. \quad (12)$$

Proof: Please see Appendix C. ■

C. Max-Min Fair Regime

Unlike the predefined QoS constraints in the previous optimization problem, the max-min fair regime aims at maximizing the minimum data rate achievable by both users across the DL and UL directions. Therefore, the max-min fair regime problem can be formulated as

$$\max_{\alpha} \left(\min_{i \in \{1,2\}} \{\hat{\gamma}_i(\alpha), \check{\gamma}_i(\alpha)\} \right) \text{ s.t. } \alpha_1 + \alpha_2 \leq 1. \quad (13)$$

Following the standard definition of max-min fairness, the worst-performing transmission should be pushed to its upper limit to enhance the overall system's performance, yielding a unique optimal max-min fair rate common for all users in both directions. It is worth noting that the optimal max-min rate (MMR) solution requires the exploitation of entire RIS elements (i.e., $\alpha_1 + \alpha_2 = 1$), thus reducing the RIS partitioning variables to a single variable (i.e., $\alpha_1 = \alpha, \alpha_2 = 1 - \alpha$). Also denoting this common rate as an auxiliary variable R_{th} , the problem in (13) can be equivalently formulated as follows

$$\begin{aligned} \mathbf{P}_3 : \max_{\alpha, R_{th}} \quad & R_{th} \\ \text{C}_3^1 : \text{ s.t. } \quad & \hat{\gamma}_i(\alpha) \geq \gamma_{th}, i \in \{1, 2\}, \\ & \check{\gamma}_i(\alpha) \geq \gamma_{th}, i \in \{1, 2\}, \\ \text{C}_3^3 : \quad & \alpha \in [0, 1], \end{aligned} \quad (14)$$

where $\gamma_{th} = 2^{R_{th}/B} - 1$. The optimal RIS portions and max-min fair rates can be obtained by solving the set of equations dictated by C_3^1 and C_3^2 , i.e., $\hat{\gamma}_1(\alpha) = \hat{\gamma}_2(\alpha) = \check{\gamma}_1(\alpha) = \check{\gamma}_2(\alpha) = R_{th}$. Following the similar solution methodology presented for the feasible region, the optimal portion α^* will be attained when the lower and upper bounds are equal to each other. However, reaching the max-min fairness achievable by the network's true capability is not possible without power control. This issue is depicted in Fig. 2, which shows the upper and lower bounds of α with respect to increasing the max-min fair rate, R_{th} . Without power control, the UB curve drops suddenly at 2.32 [bits/s/Hz] for $\beta_2 = 0.8$. This premature drop can be alleviated by introducing DL power control at BS for two reasons: 1) BS already coordinates with the RIS controller and has a higher degree of power control freedom thanks to the available high transmit power, and 2) requiring users to perform UL power control voids the spirit of inherent GF-NOMA of the proposed system model. Accordingly, we restate the previous problem with the DL power control variables as follows

$$\begin{aligned} \mathbf{P}_4 : \max_{\alpha, \beta, R_{th}} \quad & R_{th} \\ \text{C}_4^1 : \text{ s.t. } \quad & \hat{\gamma}_i(\alpha) \geq \gamma_{th}, i \in \{1, 2\}, \\ & \check{\gamma}_i(\alpha, \beta) \geq \gamma_{th}, i \in \{1, 2\}, \\ \text{C}_4^3 : \quad & \alpha \in [0, 1], \beta \in [0, 1], \end{aligned} \quad (15)$$

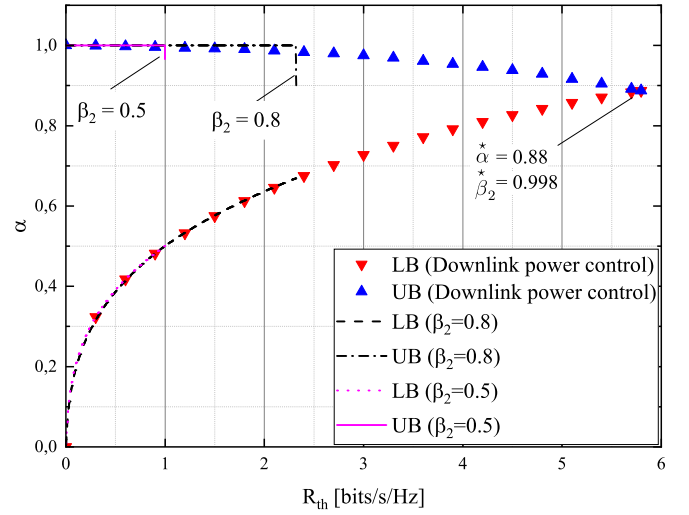


Fig. 2. The max-min fairness optimization with and without power control at BS.

where we again have a single DL power control variable since BS must exploit its allowed maximum transmit power to reach the highest possible max-min rate (i.e., $\beta_1 + \beta_2 = 1, \beta_1 = \beta, \beta_2 = 1 - \beta$). We solve this joint DL power control and RIS partitioning problem following the same intuition explained in Section III-B; we obtain optimal RIS partition and BS power allocation in two stages. The following lemma first finds the optimal portion, α^* , yielding both users' max-min fair UL rates.

Lemma 4: The optimal RIS portion that yields the max-min fair UL rate is given by

$$\alpha^* = -\frac{b_m}{4a_m} + \frac{\pm_t \bar{E}_2 \pm_s \sqrt{-[3\bar{A}_1 + 2y_0 \pm_t \frac{2\bar{B}_1}{\bar{E}_2}]}}{2}, \quad (16)$$

where both $\pm_t$ have the same sign, while $\pm_s$ is independent.

Proof: Please see the proofs and notations in Appendix D. ■ Based on given RIS portions in Lemma 4, the following lemma provides the CF DL power allocation that yields the DL rates the same as UL rates.

Lemma 5: Based on the RIS portion defined in (16), the power allocation factor that gives max-min fair DL rates equal to the UL max-min fair rates is given by

$$\beta^* = \frac{-\tilde{B}_2 \pm \sqrt{\tilde{B}_2^2 - 4\tilde{A}_2\tilde{C}_2}}{2\tilde{A}_2}. \quad (17)$$

Proof: Please see Appendix E. ■

In order to illustrate the impact of DL power control, Fig. 2 shows UB and LB of α with respect to the max-min rate auxiliary variable R_{th} . As we increase R_{th} , UB experiences a sharp drop around $R_{th} = 1$ and $R_{th} = 2.35$ for $\beta_2 = 0.5$ and $\beta_2 = 0.8$, respectively. On the other hand, the optimal BS power allocation ($\beta^* = 0.002$) allows UB and LB of α to finally match around $\alpha^* = 0.88$, yielding overall bidirectional max-min SE of $R_{th} = 5.6$ [bits/s/Hz]. That is, the lack of BS power control leads to the immature drop of the UB curve, yielding a suboptimal max-min rate performance. Fig. 2 also provides an explanation of the previous subsections' strategy

TABLE I
TABLE OF PARAMETERS

Reflection coefficient	$\rho = 1$
RIS elements	$K = \{96, 128, 256, 512\}$
BS-RIS distance	$d_{k,b} = 150$ m
RIS- U_1 distance	$d_{1,k} = 59$ m
RIS- U_2 distance	$d_{2,k} = 61$ m
Pass loss	$\tau = 2.2$ [6]
m parameter	$m = \{1, 2, 3\}$
Noise power	$\sigma_p^2 = -60$ dBm [47], [48]
Transmit power from BS	$P_b = 30$ dBm
Transmit power from U_i	$P_1 = P_2 = 10$ dBm
Rate QoS constraint	$R_{th} = \hat{R}_{th} = \check{R}_{th} = 2$ bits/s/Hz
Weight variable	$\omega = 0.5$

of finding RIS partitions first and then adjusting the BS power allocations accordingly.

D. Maximum Throughput Regime

In the maximum throughput regime, our goal is to maximize the weighted sum-rate of bidirectional traffic subject to the UL and DL QoS constraints. Hence, this problem is formulated as follows

$$\begin{aligned}
 \mathbf{P}_5 : \max_{\alpha, \beta} & \omega \sum_i \hat{\gamma}_i(\alpha) + (1 - \omega) \sum_i \check{\gamma}_i(\alpha, \beta) \\
 \text{C}_5^1 : \text{s.t. } & \hat{\gamma}_i(\alpha) \geq \hat{\gamma}_{th}^i, i \in \{1, 2\}, \\
 & \text{C}_5^2 : \check{\gamma}_i(\alpha, \beta) \geq \check{\gamma}_{th}^i, i \in \{1, 2\}, \\
 & \text{C}_5^3 : \alpha \in [0, 1], \beta \in [0, 1],
 \end{aligned} \quad (18)$$

where $\hat{R}_i = \log_2[1 + \hat{\gamma}_i(\alpha)]$ and $\check{R}_i = \log_2[1 + \check{\gamma}_i(\alpha)]$, with $\hat{R}_i$ and $\check{R}_i$ denoting the rates of U_i for the UL and DL transmissions, respectively; $0 \leq \omega \leq 1$ is the design parameter and should be set based on the underlying network's directional traffic volume; for example, symmetrical traffic would require $\omega = 0.5$. Similar to the max-min fair problems, $\mathbf{P}_5$ also has a single RIS partitioning variable since all elements must be used to reach maximum throughput (i.e., $\alpha_1 + \alpha_2 = 1$, $\alpha_1 = \alpha, \alpha_2 = 1 - \alpha$). In the next section, we solve $\mathbf{P}_5$ using metaheuristic algorithms by setting the RIS portions to their feasible regions given in Lemma 1.

IV. NUMERICAL RESULTS AND DISCUSSION

In this section, we validate the derived CF solutions and assess the performance of bidirectional NOMA traffic under various system parameters. The default simulation parameters are listed in Table I, unless specified otherwise. The problems discussed in the preceding section can be transformed into geometric programming (GP) problems [41], [42], [43], [44], [45], which can be efficiently solved using numerical solvers, such as the GP toolbox of CVX, a sophisticated convex programming toolbox designed for Matlab [46]. However, the CF derivations offered in this work are critical to achieve optimal RIS partitioning while minimizing computational complexity, hardware costs, and power consumption. Therefore, we utilize the GP solver of CVX to validate our analytical results in the following discussion. It is important to note that the

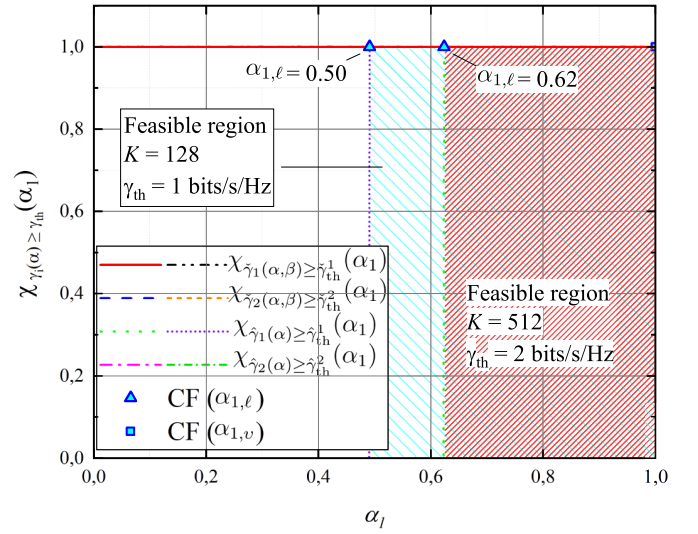


Fig. 3. Validation of derived feasibility regions.

complex-valued terms in the SINR expressions arising from the non-coherently aligned portion of RIS prevent us from transforming the problems into GP problems. As a result, the solutions obtained using CVX ignore the signals received from the non-coherently aligned portion. Despite ignoring the impact of misaligned channels throughout this section, we will demonstrate that the impact of non-coherent phases on the system performance is negligible. In the following discussion, we provide validation of our analytical findings and showcase the system's performance under various operational regimes.

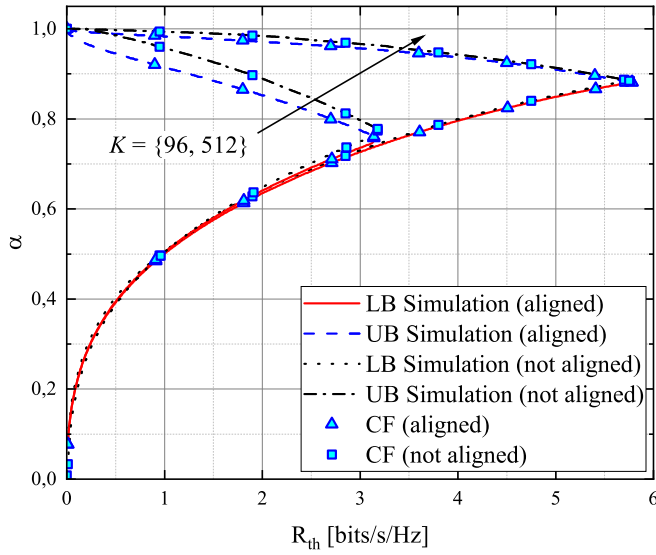
A. QoS Sufficient Regime

This subsection provides the simulation results to validate the analytically derived solutions for the QoS sufficient regime provided in Lemma 1. To this aim, we define the following indicator function to measure the feasibility of the problem as

$$\chi_{\mathcal{A}}(x) = \begin{cases} 1, & \text{if } x \in \mathcal{A}, \\ 0, & \text{otherwise.} \end{cases} \quad (19)$$

Fig. 3 illustrates the system's feasible region under different parameter setups. The vertical axis in the figure defines the satisfaction of the conditions C_0^1 and C_0^2 provided in (7) ($\chi_{\hat{\gamma}_i(\alpha) \geq \hat{\gamma}_{th}^i}(\alpha_1)$ and $\chi_{\check{\gamma}_i(\alpha) \geq \check{\gamma}_{th}^i}(\alpha_1)$).

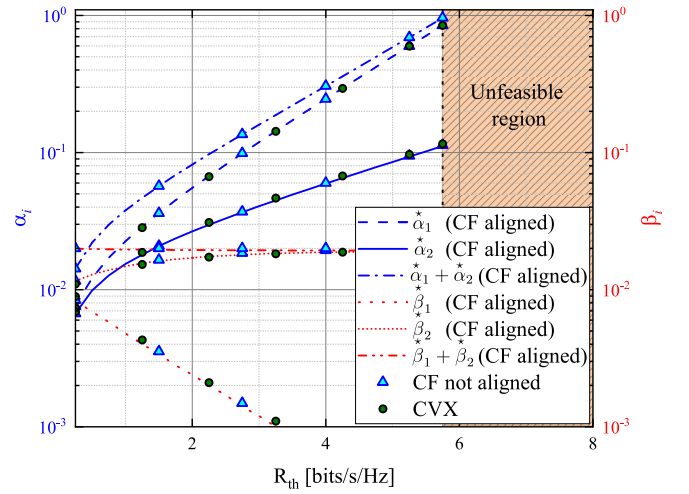
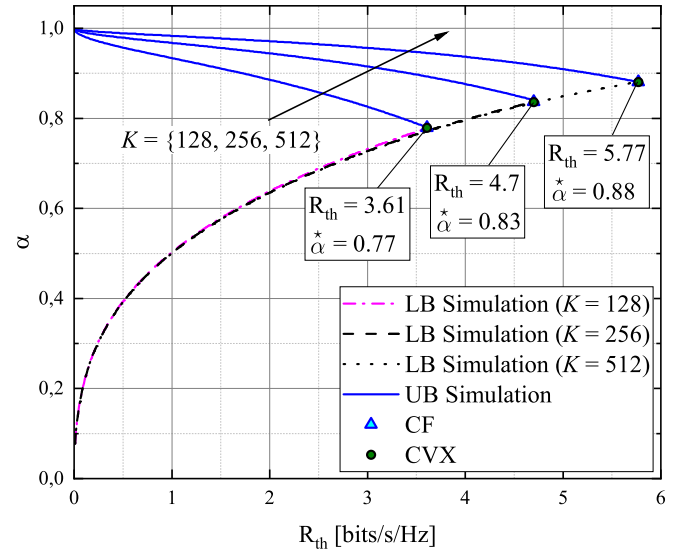
The feasibility region in this study refers to the region in which all users' SINRs surpass the desired QoS threshold value. If the user's SINR surpasses the designated threshold value, the condition is satisfied, and the logical one is assigned, signifying that the user is within the feasible region. Conversely, if the condition is violated, a zero is assigned to indicate a non-feasible region. For the system under consideration, two scenarios are set: $K = 128$ with a QoS threshold of $\gamma_{th} = 1$ bits/s/Hz, and $K = 512$ with a QoS threshold of $\gamma_{th} = 2$ bits/s/Hz. In the latter scenario, the feasible region is defined when the value of α_1 lies between 0.62 and 1, while in the former scenario, the feasibility region is broader, stretching from 0.5 to 1, due to the lower QoS demands. It is noteworthy that the simulated results align perfectly with the

Fig. 4. QoS sufficient regime for different K .

analytical performance of the feasible region, as determined by UB and LB, respectively, which characterize the range of α_1 for meeting the QoS constraints. The feasibility region in the proposed system is characterized by the range of values of α_1 for which all users' SINRs exceed the specified QoS threshold. The feasibility region becomes narrower as the rate threshold increases, until the lower and upper bounds converge at a specific point that defines the optimal RIS partitioning coefficient for achieving the maximum min-rate for both DL and UL transmissions. Fig. 4 provides a visual representation of the feasible region as a function of the rate threshold and α_1 for different K and RIS configurations. The results show that the proposed analytical solutions for the upper and lower bounds, as presented in Appendix A, accurately match the simulated performance, providing further validation of the derivations. Additionally, a comparison between the aligned and not aligned phase scenarios reveals that the difference in performance is negligible for a high number of RIS elements, justifying the assumption made in the analysis.

B. RIS and Power Efficient Regime

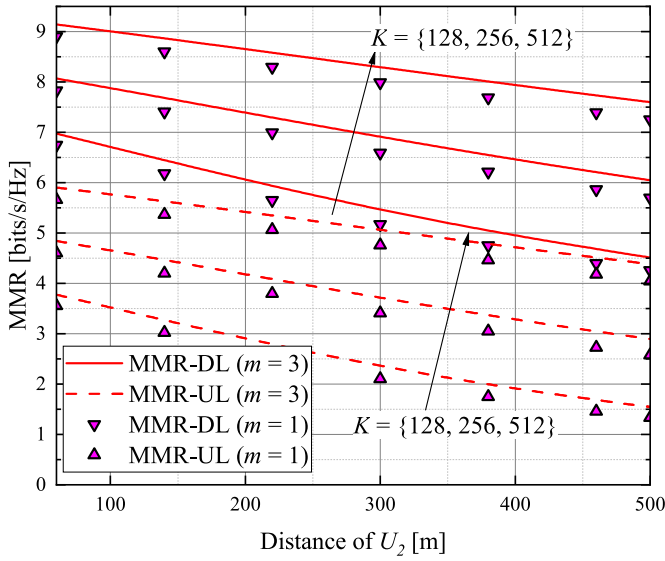
Fig. 5 presents a visual representation of the results from the optimization of RIS and power efficiency, taking into account the threshold rate, and utilizing the analytically derived coefficients α_i^* and β_i^* . The feasibility of the system is defined by C_1^2 in (9). The system reaches an infeasible state when the sum of the RIS portions surpasses the available resources, i.e., $\alpha_1^* + \alpha_2^* = 1$. It can be observed that the highest achievable rate threshold is attained at $R_{th} = 5.75$ bits/s/Hz. The findings of this analysis are in agreement with the numerical solutions obtained using the CVX toolbox, further validating the accuracy of the derived CF solutions. For the DL scenario, it is evident that the values of β_i are significantly low. This is due to the higher transmit power used in the DL scenario, $P_b = 30$ dBm, compared to the UL scenario, $P_i = 10$ dBm. Thus, the DL users' QoS conditions are easily satisfied with low

Fig. 5. Optimal α and β values for RIS and power efficient regime when $K = 512$.Fig. 6. Max-min fairness for different K .

values of β_i . The conclusion is that the DL transmission can support higher rate threshold constraints if the transmit power of users U_i is increased. It is worth mentioning that the impact of the non-coherently aligned phases on the performance is minimal, and thus, we only consider the effect of aligned channels in the subsequent analysis.

C. Max-Min Fair Regime

Fig. 6 shows the max-min fairness performance for the bidirectional transmission scenario. The CF solution for α^* , as derived in Lemma 5, is in close agreement with both CVX and simulated solutions, providing evidence for the correctness of the analytical derivations. It can be observed that MMR increases with the number of RIS elements. For instance, when the number of RIS elements increases from 128 to 512, a 47% MMR improvement is achieved. Correspondingly, the value of α^* increases from 0.77 to 0.88. It is worth noting that the improvement in performance with an increasing number of

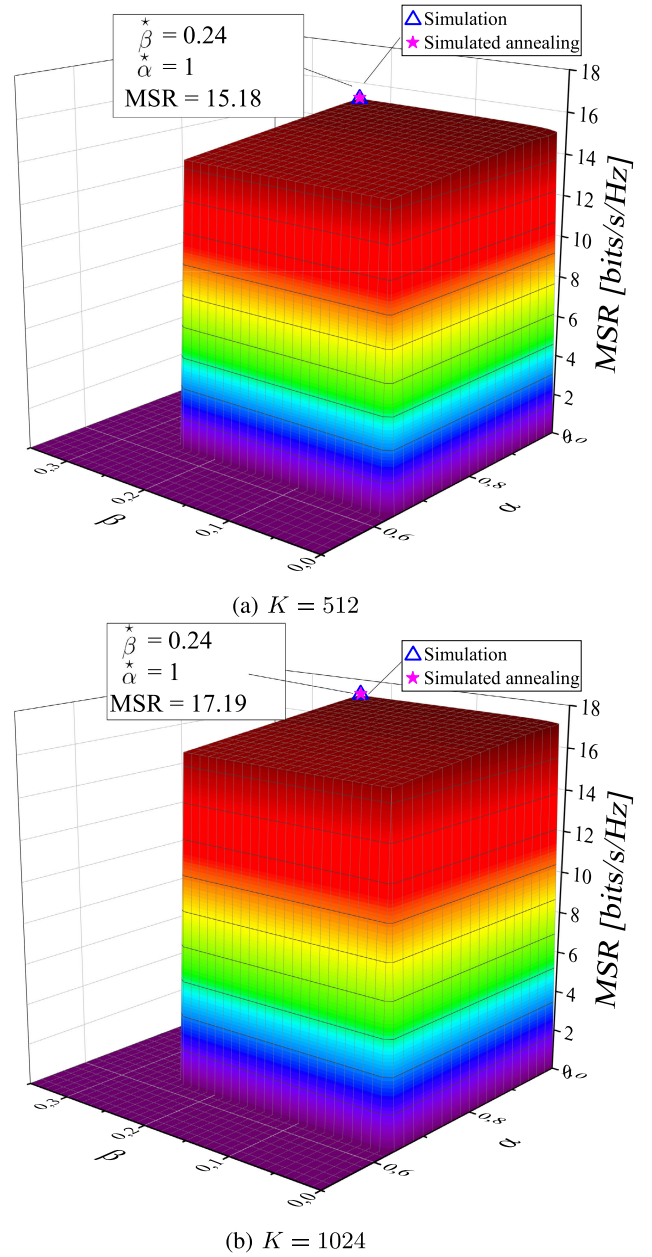
Fig. 7. MMR for different K and m .

RIS elements mainly impacts the UB performance. This is due to the impact of U_1 's SINR, which determines the LB values, in UL transmission. The signal of U_1 experiences interference from U_2 , resulting in a similar signal gain regardless of the number of RIS elements.

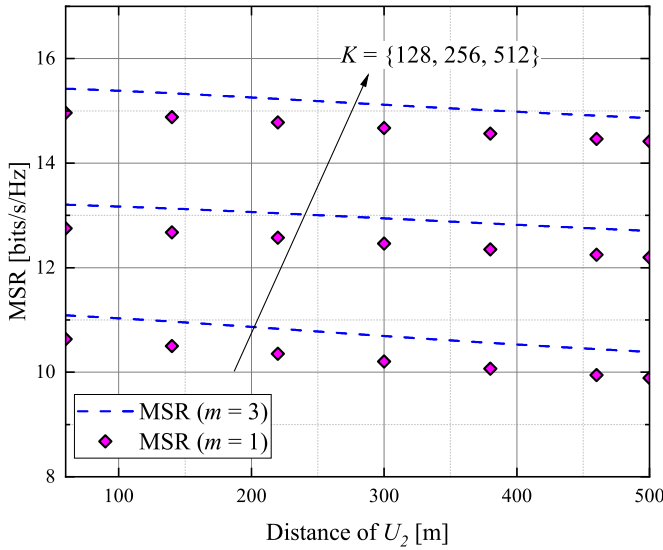
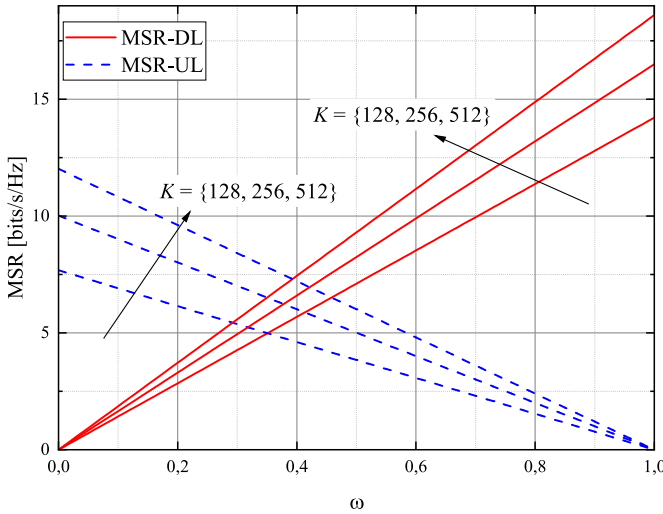
Fig. 7 presents the MMR results of the users, taking into account various system parameters and different distances of U_2 from RIS. The optimization is carried out iteratively for each new position of the NOMA user. The plot shows that as U_2 moves further away from RIS, the performance of the system deteriorates due to the increased signal loss. Additionally, the plot demonstrates that increasing the number of reflective elements results in a linear improvement in system performance. For instance, when U_2 is positioned 300m away and the Nakagami- m parameter is equal to 3, MMR-DL improves from 6.7 to 7.8 bits/s/Hz when the number of RIS elements increases from 128 to 256. Furthermore, it further enhances to 9 bits/s/Hz when $K = 512$. The plot also reveals that the m parameter of the Nakagami- m distribution has a significant impact on the MMR performance. With a lower m value of 1, the channel experiences Rayleigh fading, which leads to a lower MMR performance. This fading model is well-suited for communication with no line of sight. Conversely, increasing the value of m results in an improvement in the MMR performance as it represents the number of multi-path components in the channel, and a higher m value corresponds to a higher number of multi-paths, providing more diversity.

D. Maximum Throughput Regime

The conjoined effect of power and RIS elements allocation on the total system's MSR performance is depicted through a 3D plot shown in Fig. 8. Two scenarios were considered for the considered MSR analysis, $K = 512$ and $K = 1024$. Both the number of RIS elements and power allocation coefficients have a significant impact on the system's performance. It is observed that the system's performance is improved with

Fig. 8. The system's MSR vs. α vs. β .

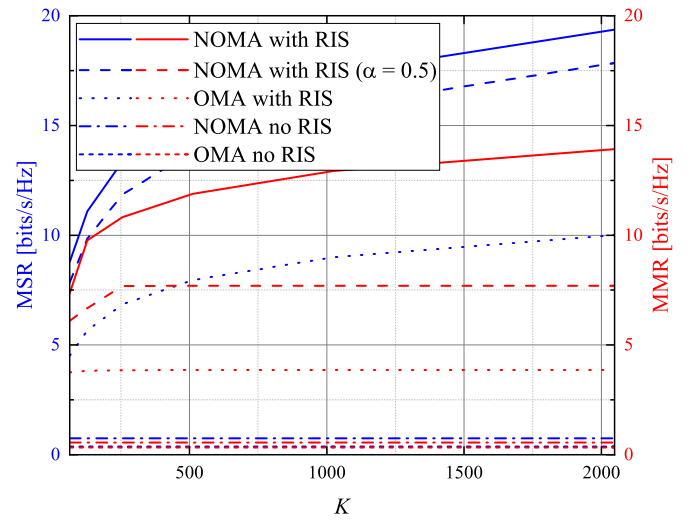
higher values of K . For instance, when $K = 1024$, the system achieves a maximum MSR of 17.19 bits/s/Hz, while the maximum MSR when $K = 512$ is 15.18 bits/s/Hz. The optimal allocation portions to attain the maximum performance for both settings are $\alpha^* = 1$ and $\beta^* = 0.24$. The optimal values were obtained using simulated annealing, a metaheuristic approach to tackle global optimization problems with continuous or discrete search space [49], [50]. Simulated annealing finds approximation to the global optimum of a given function by using probabilistic techniques to reduce the time complexity compared to the exhaustive search approach. The simulated annealing provided optimal α and β values with a total convergence time of 0.69 and 0.71 seconds for $K = 512$ and $K = 1024$, respectively. It should be noted that no MSR is

Fig. 9. MSR for different K and m .Fig. 10. MSR (DL-UL) for different K .

recorded until α reaches 0.64 and β exceeds 0.24 due to the C_5^1 and C_5^2 QoS conditions.

Fig. 9 illustrates the system's MSR performance as a function of the distance between U_2 and RIS. The optimal values of α^* and β^* are used in this analysis. As the distance between U_2 and RIS increases, the MSR performance deteriorates, but adding more RIS elements improves the system's overall throughput. The difference in U_2 positions has a minimal effect on the MSR performance due to the DL power gain, as evident from the significantly higher MMR results in Fig. 7 compared to the UL counterpart.

In Fig. 10, the system's MSR performance is compared for the DL and UL transmission scenarios. The optimal power and elements allocation coefficients are used for each ω value, resulting in a linear behavior in the obtained MSR results since ω determines the priority of the transmission scenario. As ω increases, the DL MSR improves, but the UL MSR decreases. Adding more RIS elements also enhances the

Fig. 11. MSR (NOMA-OMA) for different K .

overall system's MSR performance. For example, a nearly 27% gain in the DL MSR can be observed between $K = 128$ and $K = 512$ (from 14.21 to 16.49 bits/s/Hz).

Fig. 11 presents the comparison of MSR and MMR for NOMA and OMA systems as a function of K . The optimal partitioning and power splitting coefficients are selected based on the weight value for each scheme followed by the proposed algorithms in the corresponding optimization problems. In the conventional RIS-aided NOMA benchmark, the users are allocated with an equal number of reflective elements (i.e. $\alpha = 0.5$) without consideration of the non-aligned part of the signal. Compared to no optimal partitioning case (i.e. $\alpha = 0.5$), the proposed model is shown to provide 18.6%–57.61% and 11.04%–8.18% more MMR and MSR performance between $K = 64$ and $K = 2048$, respectively. In the OMA benchmark, all users equally share the available bandwidth; however, the corresponding power and RIS splits are optimized following the proposed procedures. The findings indicate that NOMA consistently surpasses OMA in terms of the relevant metrics, using the same simulation parameters.

For example, in the case of the system's MSR and MMR with a value of $K = 64$, and considering the RIS-assisted scenario, OMA achieves 4.52 bits/s/Hz and 3.74 bits/s/Hz, whereas NOMA achieves significantly higher rates of 8.78 bits/s/Hz and 7.35 bits/s/Hz, nearly doubling the performance in terms of both MSR and MMR. Furthermore, the outcomes demonstrate that increasing the number of RIS elements improves the performance of both NOMA and OMA networks. It is worth mentioning that the OMA and NOMA scenarios without RIS are not capable of meeting the QoS requirement of 2 bits/s/Hz for the entire range of K . These results demonstrate the superiority of the proposed RIS partitioning method for NOMA over the other benchmark scenarios.

E. Convergence Time Comparison

The convergence time comparison is presented in Fig. 12, where the left subplot compares the convergence time of

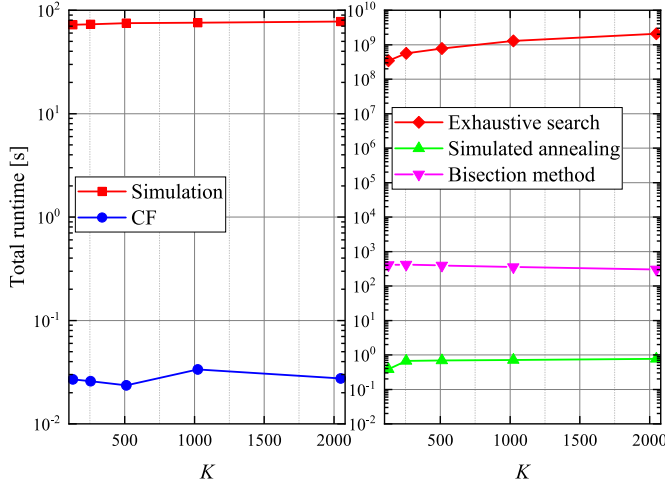
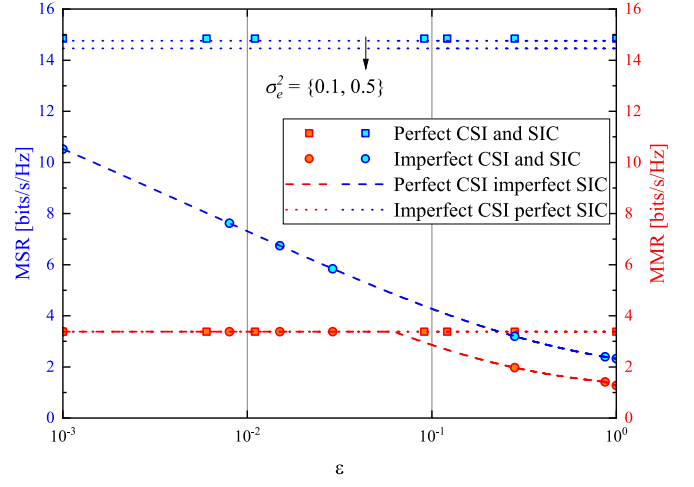


Fig. 12. Running time for different optimization problems.

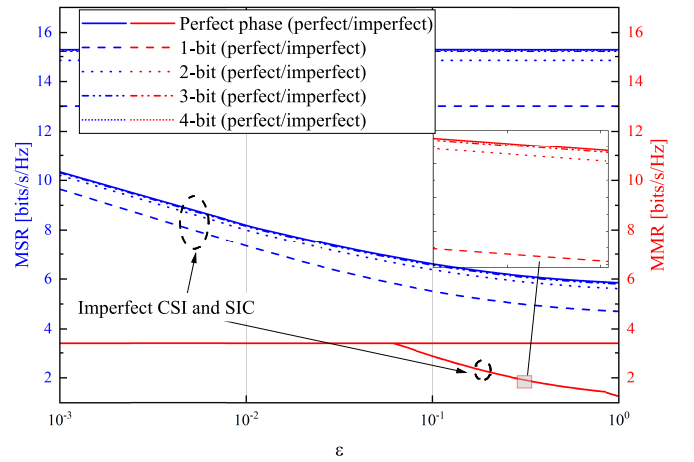
optimal CVX solutions and the proposed CF method for the QoS sufficient regime. The figure demonstrates a slight increase in convergence time of CVX solutions as the number of elements increases, which is mainly because of increasing complexity of matrix operations. More importantly, it is obvious from Fig. 12 that proposed CF solutions reduce the time spent for optimal RIS partitioning more than 3 orders of magnitude. As shown in the right sub-plot, the proposed simulated annealing approach was able to reach a sub-second convergence duration to obtain the optimal values of $\hat{\alpha}^*$ and $\hat{\beta}^*$ for the maximum throughput regime. In the right sub-plot, we compare simulated annealing approach with two benchmarks: 1) exhaustive solutions search α and β with 10^{-6} step size in brute-force approach and 2) a nested bi-section method where a bi-section perform a line search for optimal α while performing another bi-section to find the optimal β corresponding to the α value generated at each iteration. For the sake of fairness, we set tolerance of bi-section and simulated annealing approaches to 10^{-6} as well. Simulated annealing is able to return $\hat{\alpha}^*$ and $\hat{\beta}^*$ more than 2 and 8 orders of magnitude than nested bi-section and exhaustive search approaches, respectively.

F. The Impact of CSI, SIC, and Phase Shift Imperfections

The Fig. 13a depicts the effect of imperfect CSI and SIC on the maximum throughput and MMR results with ϵ and σ_e^2 denoting the imperfect SIC and CSI coefficients, respectively. The imperfection models and related values were derived from the relevant literature [51], [52], [53]. As illustrated in the figure, imperfect SIC has a greater effect on the system's performance compared to imperfect CSI. The influence of imperfect CSI is almost imperceptible for MMR results, while its effect becomes noticeable for MSR results. With the increasing ϵ , the MMR curve follows a constant rate until reaching the value of $\epsilon = 0.065$, where the MMR starts to drop. This is due to the max-min fair regime's notion of selecting the weakest SINR for further maximization described in (13). Until $\epsilon = 0.065$, the weakest SINRs were the ones that



(a) MSR for perfect/imperfect CSI/SIC.



(b) Different phase shift models comparison.

Fig. 13. Comparison to imperfect system configurations.

treated other signal messages as interference. Once ϵ exceeds the threshold value, the weakest SINR becomes the one that processes SIC.

Figure 13b compares the MSR and MMR of a system with a perfect phase shift to one with practical discrete phase shifters under CSI and SIC imperfections. Similar to Fig. 13a, the imperfect SIC and CSI has a greater impact on the system's MSR. The figure clearly shows that when a single RIS element is set to 2 bits with four phase shift options, the result closely mirrors that of the ideal scenario. Furthermore, if we select 4 bits with 16 phase shift options, the system's performance converges closely toward the ideal scenario.

V. CONCLUSION

This research examines the impact of RISs on bidirectional GF-NOMA networks. We proposed a technique that divides the RISs to optimize the channel conditions for NOMA users, resulting in enhanced NOMA gain and removing the need for power control in the uplink. The proposed method is thoroughly evaluated under four practical scenarios, including the QoS sufficient, RIS and power efficient, max-min fair, and maximum throughput regimes, each subject to both uplink

and downlink QoS constraints. We derived the closed-form solutions that show how optimal RIS partitioning can fulfil the uplink QoS requirements and optimal BS power control can ensure the downlink QoS compliance by taking advantage of the decoupled nature of the RIS portions and BS transmit power. The results of the study have been verified through simulations, highlighting the substantial benefits that RISs can bring to GF-NOMA networks in these operational scenarios. Furthermore, the superiority of the proposed GF-NOMA scheme was shown compared to other benchmark schemes. To further facilitate the practicability of the proposed schemes, future work will focus on the user pairing and RIS assignment concept to facilitate a large number of NOMA users over multiple RISs distributed over the cell area and analyze the impact of imperfect SIC/CSI transmission environment on the system's performance and complexity.

APPENDIX A PROOF OF LEMMA 1

The feasible region has LB and UB ($\alpha_{i,v}^i$ and $\alpha_{i,\ell}^i$) which are characterized by the critical values of obtained α_i^i . For deriving the feasible region of the individual user for a specific transmission scenario, each of the objective functions needs to be solved considering the constraint C_0^3 in (7). Therefore, substituting (3) into (7), the inequality equation for U_1 can be rewritten as

$$\frac{\beta_1 P_b \left\{ [\tilde{\alpha}_1^1 A_1 + \tilde{\alpha}_1^1 \Lambda_1]^2 + [\tilde{\alpha}_1^1 \lambda_1]^2 \right\}}{\sigma_1^2} \geq \gamma_{th}, \quad (A.1)$$

where Λ_i and λ_i are the real and complex parts of $\Psi_i e^{j\tilde{\theta}_{r,k}}$ in (3), respectively. Following some algebraic manipulations, we obtain the quadratic inequality as

$$[\tilde{\alpha}_1^1]^2 \tilde{A}_1 + \tilde{\alpha}_1^1 \tilde{B}_1 + \tilde{C}_1 \geq 0, \quad (A.2)$$

where $\tilde{A}_1 = A_1^2 - 2A_1\Lambda_1 + \Lambda_1^2$, $\tilde{B}_1 = 2A_1\Lambda_1 - 2\Lambda_1^2 - 2\lambda_1^2$, $\tilde{C}_1 = \Lambda_1^2 + \lambda_1^2 - \frac{\gamma_{th}\sigma_1^2}{\beta_1 P_b}$. Consequently, the roots of the quadratic inequality can be found as

$$\tilde{\alpha}_{1,\{v,\ell\}}^1 = \frac{-\tilde{B}_1 \pm \sqrt{\tilde{B}_1^2 - 4\tilde{A}_1\tilde{C}_1}}{2\tilde{A}_1}. \quad (A.3)$$

Next, in a similar manner, the inequality for U_2 for the DL transmission can be expressed as

$$\frac{\beta_2 P_b \left\{ [\tilde{\alpha}_1^2 A_2 + \tilde{\alpha}_2^2 \Lambda_2]^2 + [\tilde{\alpha}_2^2 \lambda_2]^2 \right\}}{\beta_1 P_b \left\{ [\tilde{\alpha}_1^2 A_2 + \tilde{\alpha}_2^2 \Lambda_2]^2 + [\tilde{\alpha}_2^2 \lambda_2]^2 \right\} + \sigma_2^2} \geq \gamma_{th}, \quad (A.4)$$

where $\tilde{\alpha}_1^2 = 1 - \tilde{\alpha}_2^2$. Similarly to (A.2), the quadratic inequality is defined as

$$[\tilde{\alpha}_1^2]^2 \tilde{A}_2 + \tilde{\alpha}_1^2 \tilde{B}_2 + \tilde{C}_2 \geq 0, \quad (A.5)$$

where $\tilde{A}_2 = [A_2^2 + \Lambda_2^2 + \lambda_2^2 - 2A_2\Lambda_2][\gamma_{th}\beta_1 P_b - \beta_2 P_b]$, $\tilde{B}_2 = [2A_2\Lambda_2 - 2\Lambda_2^2][\gamma_{th}\beta_1 P_b - \beta_2 P_b]$, and $[\gamma_{th}\beta_1 P_b A_2^2 + \gamma_{th}\sigma_2^2 - \beta_2 P_b A_2^2]$. Then, we obtain the roots from (A.5) as follows

$$\tilde{\alpha}_{1,\{v,\ell\}}^2 = \frac{-\tilde{B}_2 \pm \sqrt{\tilde{B}_2^2 - 4\tilde{A}_2\tilde{C}_2}}{2\tilde{A}_2}, \quad (A.6)$$

Next, inserting (6) into (7), the QoS condition for U_1 in the UL transmission is written as

$$\frac{P_1 \left\{ [\hat{\alpha}_1^1 B_1 + \hat{\alpha}_2^1 \Omega_1]^2 + [\hat{\alpha}_2^1 \rho_1]^2 \right\}}{P_2 \left\{ [\hat{\alpha}_2^1 B_2 + \hat{\alpha}_1^1 \Omega_2]^2 + [\hat{\alpha}_1^1 \rho_2]^2 \right\} + \sigma_0^2} \geq \gamma_{th}, \quad (A.7)$$

where $\hat{\alpha}_2^1 = 1 - \hat{\alpha}_1^1$; Ω_i and ρ_i are the real and complex parts of $\Upsilon_i e^{j\hat{\theta}_{r,k}}$ in (6). Then, we re-express the inequality as follows

$$[\hat{\alpha}_1^1]^2 \hat{A}_1 + \hat{\alpha}_1^1 \hat{B}_1 + \hat{C}_1 \geq 0, \quad (A.8)$$

where $\hat{A}_1 = \gamma_{th} P_2 [B_2^2 + \Omega_2^2 - 2B_2\Omega_2 + \rho_2^2] - P_1 [B_1^2 + \Omega_1^2 - 2B_1\Omega_1 + \rho_1^2]$, $\hat{B}_1 = \gamma_{th} P_2 [2B_2\Omega_2 - \Omega_2^2] - P_1 [2B_1\Omega_1 - 2\Omega_1^2 - 2\rho_1^2]$, and $\hat{C}_1 = \gamma_{th} [P_2 B_2^2 + \sigma_b^2] - P_1 [\Omega_1^2 + \rho_1^2]$. The lower and upper bounds for UL U_1 can be found as

$$\hat{\alpha}_{1,\{v,\ell\}}^1 = \frac{-\hat{B}_1 \pm \sqrt{\hat{B}_1^2 - 4\hat{A}_1\hat{C}_1}}{2\hat{A}_1}. \quad (A.9)$$

Now, the condition for U_2 can be written as

$$\frac{P_2 \left\{ [\hat{\alpha}_2^2 B_2 + \hat{\alpha}_1^2 \Omega_2]^2 + [\hat{\alpha}_1^2 \rho_2]^2 \right\}}{\sigma_0^2} \geq \gamma_{th}. \quad (A.10)$$

Then, after some mathematical manipulations, (A.10) can be re-written as

$$[\hat{\alpha}_1^2]^2 \hat{A}_2 + \hat{\alpha}_1^2 \hat{B}_2 + \hat{C}_2 \geq 0, \quad (A.11)$$

where $\hat{\alpha}_2^2 = 1 - \hat{\alpha}_1^2$, $\hat{A}_2 = B_2^2 + \Omega_2^2 - 2B_2\Omega_2 + \rho_2^2$, $\hat{B}_2 = 2B_2\Omega_2 - B_2^2$, and $\hat{C}_2 = B_2^2 - \frac{\gamma_{th}\sigma_b^2}{P_2}$. The final roots are written as

$$\hat{\alpha}_{1,\{v,\ell\}}^2 = \frac{-\hat{B}_2 \pm \sqrt{\hat{B}_2^2 - 4\hat{A}_2\hat{C}_2}}{2\hat{A}_2}. \quad (A.12)$$

APPENDIX B PROOF OF LEMMA 2

The critical values of the elements' allocation coefficient for each user in both DL and UL transmission scenarios can be derived with the assistance of the C_1^1 and C_1^2 conditions in (9). Firstly, the minimum values of allocation coefficients are found for the UL transmission scenario. Then, the values of α_i^* are used for defining the minimum power allocation parameters in the DL transmission. For deriving the optimal elements allocation of the individual user for the UL scenario, each of the objective functions needs to be solved considering the constraint C_1^3 in (9).

$$\begin{cases} \hat{\gamma}_1(\alpha_1, \alpha_2) = \hat{\gamma}_{th}^1, \\ \hat{\gamma}_2(\alpha_1, \alpha_2) = \hat{\gamma}_{th}^2. \end{cases} \quad (B.1)$$

Therefore, substituting (3) into (B.1), the system of equation for UL can be rewritten as

$$\begin{cases} \frac{P_1 [\alpha_1 B_1]^2}{P_2 [\alpha_2 B_2]^2 + \sigma_b^2} = \hat{\gamma}_{th}^1, \\ \frac{P_2 [\alpha_2 B_2]^2}{\sigma_b^2} = \hat{\gamma}_{th}^2, \end{cases} \quad (B.2)$$

where $B_i = |g_{i,k}| |h_{i,k}| \psi_i$. After following the corresponding algebraic manipulations and conditions, the derived α_i^* can be expressed as in (10).

APPENDIX C PROOF OF LEMMA 3

Following the obtained values of $\hat{\alpha}_i$, we find $\hat{\beta}_i$ by inserting $\hat{\alpha}_i$ into the corresponding system of equation, likewise in (B.1)

$$\begin{cases} \frac{\beta_1 P_b [\hat{\alpha}_1 A_1]^2}{\sigma_1^2} = \tilde{\gamma}_{th}^1, \\ \frac{\beta_2 P_b [\hat{\alpha}_2 A_2]^2}{\beta_1 P_b [\hat{\alpha}_2 A_2]^2 + \sigma_2^2} = \tilde{\gamma}_{th}^2, \end{cases} \quad (C.1)$$

Similarly, following the algebraic manipulations and finding positive roots of the quadratic equation, derived $\hat{\beta}_i$ are obtained as in (12).

APPENDIX D PROOF OF LEMMA 4

The optimal element allocations for $\mathbf{P}_4$ in (15) at which the max-min rate achieved is defined as $\hat{\alpha}$ that can be obtained by equating the SINRs of both users in the UL transmission as

$$\begin{aligned} & \frac{P_1 \left\{ [\alpha B_1 + [1 - \alpha] \Omega_1]^2 + [[1 - \alpha] \rho_1]^2 \right\}}{P_2 \left\{ [[1 - \alpha] B_2 + \alpha \Omega_2]^2 + [\alpha \rho_2]^2 \right\} + \sigma_b^2} \\ &= \frac{P_2 \left\{ [[1 - \alpha] B_2 + \alpha \Omega_2]^2 + [\alpha \rho_2]^2 \right\}}{\sigma_b^2}. \end{aligned} \quad (D.1)$$

The value of the max-min threshold rate is dictated solely by the UL transmission. This is due to the fact that the UL transmission has lower power support compared to DL. Therefore, we push the weakest transmission scenario to its upper limits. After some algebraic manipulations on (D.1), the equation can be rewritten as

$$\frac{P_1 [\alpha^2 \tilde{D}_1 + 2\alpha \tilde{E}_1 + \tilde{F}_1]}{P_2 [\alpha^2 \tilde{A}_1 + 2\alpha \tilde{B}_1 + \tilde{C}_1] + \sigma_b^2} = \frac{P_2 [\alpha^2 \tilde{A}_1 + 2\alpha \tilde{B}_1 + \tilde{C}_1]}{\sigma_b^2}, \quad (D.2)$$

where $\tilde{A}_1 = B_2^2 + \Omega_2^2 + \rho_2^2 - 2B_2\Omega_2$, $\tilde{B}_1 = B_2\Omega_2 - \Omega_2^2$, $\tilde{C}_1 = \Omega_2^2$, $\tilde{D}_1 = B_1^2 + \Omega_1^2 + \rho_1^2 - 2B_1\Omega_1$, $\tilde{E}_1 = B_1\Omega_1 - \Omega_1^2 - \rho_1^2$, and $\tilde{F}_1 = \Omega_1^2 + \rho_1^2$. Finally, (D.2) is reformulated to form a polynomial equation and written as

$$a_m [\alpha]^4 + b_m [\alpha]^3 + c_m [\alpha]^2 + d_m \alpha + e_m = 0, \quad (D.3)$$

where $a_m = P_2^2 \tilde{A}_1^2$, $b_m = 4P_2^2 \tilde{A}_1 \tilde{B}_1$, $c_m = P_2^2 [2\tilde{A}_1 \tilde{C}_1 + 4\tilde{B}_1^2] + P_2 \sigma_b^2 \tilde{A}_1 - P_1 \sigma_b^2 \tilde{D}_1$, $d_m = 4P_2^2 \tilde{B}_1 \tilde{C}_1 + 2P_2 \sigma_b^2 \tilde{B}_1 - 2P_1 \sigma_b^2 \tilde{E}_1$, $e_m = P_2^2 \tilde{C}_1 + P_2 \sigma_b^2 \tilde{C}_1 - P_1 \sigma_b^2 \tilde{F}_1$.

The following fourth-degree polynomial, also called a quartic polynomial, can be solved using Ferrari's method as shown in [54] and [55, Eq. 18-27] and the final solution for the elements allocation coefficient is expressed as in (16), where both $\pm_t$ have the same sign, while $\pm_s$ is independent

and results in four roots. The rest notations are provided as follows

$$\begin{aligned} \bar{A}_2 &= -\frac{\bar{A}_1^2}{12} - \bar{C}_1, \\ \bar{B}_2 &= -\frac{\bar{A}_1^3}{108} + \frac{\bar{A}_1 \bar{C}_1}{3} - \frac{\bar{B}_1^2}{8}, \\ \bar{C}_2 &= -\frac{\bar{B}_2}{2} + \sqrt{\frac{\bar{B}_2^2}{4} + \frac{\bar{A}_2^3}{27}}, \\ \bar{D}_2 &= \bar{D}_2^{\frac{1}{3}}, \\ \bar{E}_2 &= \sqrt{\bar{A}_1 + 2y_0}, \\ \bar{A}_1 &= -\frac{3b_m^2}{8a_m^2} + \frac{c_m}{a_m}, \\ \bar{B}_1 &= \frac{b_m^3}{8a_m^3} - \frac{b_m c_m}{2a_m^2} + \frac{d_m}{a_m}, \\ \bar{C}_1 &= -\frac{3b_m^4}{256a_m^4} + \frac{c_m b_m^2}{16a_m^3} - \frac{b_m d_m}{4a_m^2} + \frac{e_m}{a_m}, \end{aligned} \quad (D.4)$$

where y_0 denotes the root of the cubic equation reduced and can be calculated as

$$y_0 = -\frac{5\bar{A}_1}{6} + \begin{cases} -\bar{B}_2^{\frac{1}{3}}, & \text{if } \bar{D}_2 = 0, \\ \bar{D}_2 - \frac{\bar{A}_2}{3\bar{D}_2}, & \text{if } \bar{D}_2 \neq 0. \end{cases} \quad (D.5)$$

This concludes the proof. It is also worth noting that, by excluding Ω_i and ρ_i from the calculations, we can obtain the optimal values of elements allocation for the case when not aligned phases are not considered.

APPENDIX E PROOF OF LEMMA 5

In order to find $\hat{\beta}$, we first insert $\hat{\alpha}$ found in Lemma 4 into the equated SINRs

$$\begin{aligned} & \frac{\beta P_b \left\{ [\hat{\alpha} A_1 + [1 - \hat{\alpha}] \Lambda_1]^2 + [[1 - \hat{\alpha}] \lambda_1]^2 \right\}}{\sigma_1^2} \\ &= \frac{[1 - \beta] P_b \left\{ [[1 - \hat{\alpha}] A_2 + \hat{\alpha} \Lambda_2]^2 + [\hat{\alpha} \lambda_2]^2 \right\}}{\beta P_b \left\{ [[1 - \hat{\alpha}] A_2 + \hat{\alpha} \Lambda_2]^2 + [\hat{\alpha} \lambda_2]^2 \right\} + \sigma_2^2}. \end{aligned} \quad (E.1)$$

Next, by solving the above equation for β , we obtain the roots of a quadratic equation as in (17), where $\hat{\beta}_2 = 1 - \hat{\beta}$, $\tilde{A}_2 = P_b^2 \rho_3 \rho_4$, $\tilde{B}_2 = P_b \sigma_2^2 \rho_3 + P_b \sigma_1^2 \rho_4$, $\tilde{C}_2 = -P_b \sigma_1^2 \rho_4$ with $\rho_3 = \left\{ \hat{\alpha} A_1 + [1 - \hat{\alpha}] \Lambda_1 \right\}^2 + \left\{ [1 - \hat{\alpha}] \lambda_1 \right\}^2$ and $\rho_4 = \left\{ [1 - \hat{\alpha}] A_2 + \hat{\alpha} \Lambda_2 \right\}^2 + \left\{ \hat{\alpha} \lambda_2 \right\}^2$.

REFERENCES

- [1] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Jan. 2020.
- [2] X. Yuan, Y. A. Zhang, Y. Shi, W. Yan, and H. Liu, "Reconfigurable-intelligent-surface empowered wireless communications: Challenges and opportunities," *IEEE Wireless Commun.*, vol. 28, no. 2, pp. 136–143, Apr. 2021.

- [3] Q. Li, M. Wen, S. Wang, G. C. Alexandropoulos, and Y.-C. Wu, "Space shift keying with reconfigurable intelligent surfaces: Phase configuration designs and performance analysis," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 322–333, 2021.
- [4] S. Zeng, H. Zhang, B. Di, Z. Han, and L. Song, "Reconfigurable intelligent surface (RIS) assisted wireless coverage extension: RIS orientation and location optimization," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 269–273, Jan. 2021.
- [5] A. A. Boulogeorgos and A. Alexiou, "Pathloss modeling of reconfigurable intelligent surface assisted THz wireless systems," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2021, pp. 1–6.
- [6] T. Hou, Y. Liu, Z. Song, X. Sun, Y. Chen, and L. Hanzo, "Reconfigurable intelligent surface aided NOMA networks," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2575–2588, Nov. 2020.
- [7] Z. Li et al., "Energy efficient reconfigurable intelligent surface enabled mobile edge computing networks with NOMA," *IEEE Trans. Cognit. Commun. Netw.*, vol. 7, no. 2, pp. 427–440, Jun. 2021.
- [8] Y. Liu, X. Mu, X. Liu, M. Di Renzo, Z. Ding, and R. Schober, "Reconfigurable intelligent surface-aided multi-user networks: Interplay between NOMA and RIS," *IEEE Wireless Commun.*, vol. 29, no. 2, pp. 169–176, Apr. 2022.
- [9] M. B. Shahab, R. Abbas, M. Shirvanimoghaddam, and S. J. Johnson, "Grant-free non-orthogonal multiple access for IoT: A survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1805–1838, 3rd Quart., 2020.
- [10] R. Abbas, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "A novel analytical framework for massive grant-free NOMA," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 2436–2449, Mar. 2019.
- [11] Y. Xiu et al., "Reconfigurable intelligent surfaces aided mmWave NOMA: Joint power allocation, phase shifts, and hybrid beamforming optimization," *IEEE Trans. Wireless Commun.*, vol. 20, no. 12, pp. 8393–8409, Dec. 2021.
- [12] C. Zhang, W. Yi, Y. Liu, K. Yang, and Z. Ding, "Reconfigurable intelligent surfaces aided multi-cell NOMA networks: A stochastic geometry model," *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 951–966, Feb. 2022.
- [13] Z. Zhang, C. Zhang, C. Jiang, F. Jia, J. Ge, and F. Gong, "Improving physical layer security for reconfigurable intelligent surface aided NOMA 6G networks," *IEEE Trans. Veh. Technol.*, vol. 70, no. 5, pp. 4451–4463, May 2021.
- [14] C. Gong, X. Yue, X. Wang, X. Dai, R. Zou, and M. Essaidi, "Intelligent reflecting surface aided secure communications for NOMA networks," *IEEE Trans. Veh. Technol.*, vol. 71, no. 3, pp. 2761–2773, Mar. 2022.
- [15] X. Tang, D. Wang, R. Zhang, Z. Chu, and Z. Han, "Jamming mitigation via aerial reconfigurable intelligent surface: Passive beamforming and deployment optimization," *IEEE Trans. Veh. Technol.*, vol. 70, no. 6, pp. 6232–6237, Jun. 2021.
- [16] M. Jung, W. Saad, M. Debbah, and C. S. Hong, "On the optimality of reconfigurable intelligent surfaces (RISs): Passive beamforming, modulation, and resource allocation," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4347–4363, Jul. 2021.
- [17] M. Fu, Y. Zhou, Y. Shi, and K. B. Letaief, "Reconfigurable intelligent surface empowered downlink non-orthogonal multiple access," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 3802–3817, Jun. 2021.
- [18] J. Zuo, Y. Liu, L. Yang, L. Song, and Y.-C. Liang, "Reconfigurable intelligent surface enhanced NOMA assisted backscatter communication system," *IEEE Trans. Veh. Technol.*, vol. 70, no. 7, pp. 7261–7266, Jul. 2021.
- [19] X. Liu, Y. Liu, Y. Chen, and H. V. Poor, "RIS enhanced massive non-orthogonal multiple access networks: Deployment and passive beamforming design," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 4, pp. 1057–1071, Apr. 2021.
- [20] Z. Yang, Y. Liu, Y. Chen, and N. Al-Dhahir, "Machine learning for user partitioning and phase shifters design in RIS-aided NOMA networks," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7414–7428, Nov. 2021.
- [21] R. Zhong, Y. Liu, X. Mu, Y. Chen, and L. Song, "AI empowered RIS-assisted NOMA networks: Deep learning or reinforcement learning?" *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, pp. 182–196, Jan. 2022.
- [22] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-lin, and Z. Wang, "Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sep. 2015.
- [23] A. Celik, "Grant-free NOMA: A low complexity power control through user clustering," TechRxiv, May 2022, doi: 10.36227/techrxiv.19688019.v1.
- [24] E. Balevi, F. T. A. Rabee, and R. D. Gitlin, "ALOHA-NOMA for massive machine-to-machine IoT communication," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–5.
- [25] C. Zhang, Y. Liu, and Z. Ding, "Semi-grant-free NOMA: A stochastic geometry model," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1197–1213, Feb. 2022.
- [26] J. Liu, G. Wu, X. Zhang, S. Fang, and S. Li, "Modeling, analysis, and optimization of grant-free NOMA in massive MTC via stochastic geometry," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4389–4402, Mar. 2021.
- [27] S. Dogan, A. Tusha, and H. Arslan, "NOMA with index modulation for uplink URLLC through grant-free access," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 6, pp. 1249–1257, Oct. 2019.
- [28] N. Ye, X. Li, H. Yu, A. Wang, W. Liu, and X. Hou, "Deep learning aided grant-free NOMA toward reliable low-latency access in tactile Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 15, no. 5, pp. 2995–3005, May 2019.
- [29] T. Ding, X. Yuan, and S. C. Liew, "Sparsity learning-based multiuser detection in grant-free massive-device multiple access," *IEEE Trans. Wireless Commun.*, vol. 18, no. 7, pp. 3569–3582, Jul. 2019.
- [30] J. Chen, L. Guo, J. Jia, J. Shang, and X. Wang, "Resource allocation for IRS assisted SGF NOMA transmission: A MADRL approach," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 4, pp. 1302–1316, Apr. 2022.
- [31] M. Fayaz, W. Yi, Y. Liu, and A. Nallanathan, "Transmit power pool design for grant-free NOMA-IoT networks via deep reinforcement learning," *IEEE Trans. Wireless Commun.*, vol. 20, no. 11, pp. 7626–7641, Nov. 2021.
- [32] A. Khaleel and E. Basar, "A novel NOMA solution with RIS partitioning," *IEEE J. Sel. Topics Signal Process.*, vol. 16, no. 1, pp. 70–81, Jan. 2022.
- [33] R. A. Tasci et al., "RIS-assisted grant-free NOMA," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2023, pp. 1–6.
- [34] E. Arslan et al., "Reconfigurable intelligent surface enabled over-the-air uplink NOMA," *IEEE Trans. Green Commun. Netw.*, vol. 7, no. 2, pp. 814–826, Jun. 2023.
- [35] M. Makin, G. Nauryzbayev, S. Arzykulov, and M. S. Hashmi, "Performance of large intelligent surface-enabled cooperative networks over Nakagami- m channels," in *Proc. IEEE 94th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2021, pp. 1–6.
- [36] A. Abdallah, A. Celik, M. M. Mansour, and A. M. Eltawil, "Deep learning-based frequency-selective channel estimation for hybrid mmWave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 3804–3821, Jun. 2022.
- [37] A. Abdallah, A. Celik, M. M. Mansour, and A. M. Eltawil, "RIS-aided mmWave MIMO channel estimation using deep learning and compressive sensing," *IEEE Trans. Wireless Commun.*, vol. 22, no. 5, pp. 3503–3521, May 2023.
- [38] X. Tan, Z. Sun, J. M. Jornet, and D. Pados, "Increasing indoor spectrum sharing capacity using smart reflect-array," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.
- [39] Z. Ding, M. Peng, and H. V. Poor, "Cooperative non-orthogonal multiple access in 5G systems," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1462–1465, Aug. 2015.
- [40] S. Arzykulov, G. Nauryzbayev, M. S. Hashmi, A. M. Eltawil, K. M. Rabie, and S. Seilov, "Hardware- and interference-limited cognitive IoT relaying NOMA networks with imperfect SIC over generalized non-homogeneous fading channels," *IEEE Access*, vol. 8, pp. 72942–72956, 2020.
- [41] A. Celik, M.-C. Tsai, R. M. Radaydeh, F. S. Al-Qahtani, and M.-S. Alouini, "Distributed cluster formation and power-bandwidth allocation for imperfect NOMA in DL-HetNets," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1677–1692, Feb. 2019.
- [42] S. Arzykulov, A. Celik, G. Nauryzbayev, and A. M. Eltawil, "UAV-assisted cooperative & cognitive NOMA: Deployment, clustering, and resource allocation," 2020, *arXiv:2008.11356*.
- [43] A. Celik, M.-C. Tsai, R. M. Radaydeh, F. S. Al-Qahtani, and M.-S. Alouini, "Distributed user clustering and resource allocation for imperfect NOMA in heterogeneous networks," *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 7211–7227, Oct. 2019.
- [44] A. Celik, R. M. Radaydeh, F. S. Al-Qahtani, A. H. A. El-Malek, and M.-S. Alouini, "Resource allocation and cluster formation for imperfect NOMA in DL/UL decoupled HetNets," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Singapore, Dec. 2017, pp. 1–6.

- [45] A. Celik, F. S. Al-Qahtani, R. M. Radaydeh, and M.-S. Alouini, "Cluster formation and joint power-bandwidth allocation for imperfect NOMA in DL-HetNets," in *Proc. IEEE Global Commun. Conf.*, Dec. 2017, pp. 1–6.
- [46] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control* (Lecture Notes in Control and Information Sciences), V. Blondel, S. Boyd, and H. Kimura, Eds. New York, NY, USA: Springer-Verlag, 2008, pp. 95–110.
- [47] M. Hua, L. Yang, Q. Wu, C. Pan, C. Li, and A. L. Swindlehurst, "UAV-assisted intelligent reflecting surface symbiotic radio system," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 5769–5785, Sep. 2021.
- [48] J. An, L. Wang, C. Xu, L. Gan, and L. Hanzo, "Optimal pilot power based channel estimation improves the throughput of intelligent reflective surface assisted systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 16202–16206, Dec. 2020.
- [49] R. A. Rutembar, "Simulated annealing algorithms: An overview," *IEEE Circuits Devices Mag.*, vol. 5, no. 1, pp. 19–26, Jan. 1989.
- [50] D. Bertsimas and J. Tsitsiklis, "Simulated annealing," *Statist. Sci.*, vol. 8, no. 1, pp. 10–15, 1993.
- [51] S. K. Singh, K. Agrawal, K. Singh, C.-P. Li, and Z. Ding, "NOMA enhanced hybrid RIS-UAV-assisted full-duplex communication system with imperfect SIC and CSI," *IEEE Trans. Commun.*, vol. 70, no. 11, pp. 7609–7627, Nov. 2022.
- [52] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "A framework of robust transmission design for IRS-aided MISO communications with imperfect cascaded channels," *IEEE Trans. Signal Process.*, vol. 68, pp. 5092–5106, 2020.
- [53] H. Zheng et al., "Robust transmission design for RIS-aided wireless communication with both imperfect CSI and transceiver hardware impairments," *IEEE Internet Things J.*, vol. 10, no. 5, pp. 4621–4635, Mar. 2023.
- [54] S. L. Shmakov, "A Universal method of solving quartic equations," *Int. J. Pure Appl. Math.*, vol. 71, no. 2, pp. 251–259, 2011.
- [55] S.-Y. Jung, J. Hong, and K. Nam, "Current minimizing torque control of the IPMSM using Ferrari's method," *IEEE Trans. Power Electron.*, vol. 28, no. 12, pp. 5603–5617, Dec. 2013.



Madi Makin (Graduate Student Member, IEEE) received the B.Sc. degree in electrical and computer engineering from Nazarbayev University, Astana, Kazakhstan, where he is currently pursuing the master's degree in electrical and computer engineering. His current research interests include wireless communication systems, with a particular focus on reconfigurable intelligent surfaces, non-orthogonal multiple access (NOMA) networks, and physical layer security.



Sultangali Arzykulov (Member, IEEE) received the B.Sc. degree (Hons.) in radio engineering, electronics, and telecommunications from Kazakh National Research Technical University, Almaty, Kazakhstan, in 2010, the M.Sc. degree in communication engineering from The University of Manchester, Manchester, U.K., in 2013, and the Ph.D. degree in science, engineering, and technology from Nazarbayev University, Astana, Kazakhstan, in 2019, where he was a Post-Doctoral Scholar from 2019 to 2020. He is currently a Post-Doctoral Fellow of the King Abdul-

lah University of Science and Technology (KAUST), Thuwal, Saudi Arabia. His research interests include wireless communication systems, with particular focus on intelligent reconfigurable surface, cognitive radio, energy harvesting, interference mitigation, and nonorthogonal multiple access. He was a technical program committee member of numerous IEEE Communication Society Flagship Conferences. He is currently a reviewer for several international journals/conferences.



Abdulkadir Celik (Senior Member, IEEE) received the M.S. degree in electrical engineering in 2013, the M.S. degree in computer engineering in 2015, and the Ph.D. degree in co-majors of electrical engineering and computer engineering from Iowa State University, Ames, IA, USA, in 2016. He was a Post-Doctoral Fellow with the King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia, from 2016 to 2020, where he is currently a Research Scientist with the Communication and Computing Systems Laboratory (CCSL).

His research interests include wireless communication systems and networks. He is currently an Editor of the IEEE WIRELESS COMMUNICATION LETTERS and an Associate Editor of *Frontiers in Communications and Networks*. He has been a technical committee member of various symposia at flagship conferences of the IEEE Communications Society and IEEE Vehicular Technology Society.



Ahmed M. Eltawil (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees (Hons.) from Cairo University, Giza, Egypt, in 1997 and 1999, respectively, and the Ph.D. degree from the University of California at Los Angeles, Los Angeles, CA, USA, in 2003. He has been a Professor in electrical engineering and computer science with the University of California at Irvine (UCI), Irvine, CA, USA, since 2005. He is currently a Professor in electrical and computer engineering with the King Abdullah University of Science and Technology (KAUST),

Thuwal, Saudi Arabia, where he joined the Division of Computer, Electrical and Mathematical Sciences and Engineering (CEMSE), in 2019. At KAUST, he is a Founder and the Director of the Communication and Computing Systems Laboratory (CCSL). His research interests include the area of efficient architectures for computing and communications systems in general and wireless systems in particular, spanning the application domains of body area networks, low-power mobile systems, machine learning platforms, sensor networks, and critical infrastructure networks. He is a Senior Member of the National Academy of Inventors, USA. He received several awards, including the NSF CAREER Award supporting his research in low-power computing and communication systems. He is a Distinguished Lecturer with the IEEE. He was selected as the Innovator of the Year of 2021 by the Henry Samueli School of Engineering, UCI. For his contributions to societal benefit through wireless innovations, he received two certificates of recognition from the United States Congress. He has been on the technical program committees and steering committees for numerous workshops, symposia, and conferences in the areas of low-power computing and wireless communication system design.



Galymzhan Nauryzbayev (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees (Hons.) in radio engineering, electronics, and telecommunications from the Almaty University of Power Engineering and Telecommunications, Almaty, Kazakhstan, in 2009 and 2011, respectively, and the Ph.D. degree in wireless communications from The University of Manchester, U.K., in 2016. From 2016 to 2018, he held several academic and research positions with Nazarbayev University, Kazakhstan, L. N. Gumi-lyov Eurasian National University, Kazakhstan, and

Hamad Bin Khalifa University, Qatar. In 2019, he joined Nazarbayev University with the rank of an Assistant Professor. His research interests include wireless communication systems, with a particular focus on reconfigurable intelligent surface-enabled communications, multiuser MIMO systems, cognitive radio, signal processing, energy harvesting, visible light communications, NOMA, and interference mitigation. He served as a technical program committee member for numerous IEEE flagship conferences. He is the Vice-Chair of the National Research Council of the Republic of Kazakhstan.