This article has been accepted for publication in IEEE Transactions on Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2022.3177773

Power Minimization for Uplink RIS-Assisted CoMP-NOMA Networks With GSIC

Hong Wang, Chen Liu, Zheng Shi, Yaru Fu, and Rongfang Song

***Abstract*—To accommodate the stringent requirements of massive connectivity and ultra-high throughput, both reconfigurable intelligent surface (RIS) and non-orthogonal multiple access (NOMA) have been perceived as the key techniques for future communication networks. In this paper, a versatile framework is conceived to boost transmit power efficiency for RIS-enabled multi-group NOMA networks in the presence of coordinated multi-point (CoMP) reception and imperfect successive interference cancellation (SIC). Particularly, a group-level SIC (GSIC) method is proposed to eliminate the decoded group's interference together with the well-designed transceivers to mitigate the aggregated interference, including the intra-group interference, the residual interference caused by imperfect SIC, and the interference of NOMA users decoded later. According to the novel framework, a power minimization problem is formulated by collaboratively optimizing the transmit power and the phase shifts. To render the problem tractable, an alternating scheme is developed to optimize the transmit power and the phase shifts iteratively. Specifically, the transmit powers for the users in the same group are devised by a parallel iteration algorithm, whilst the phase shifts are optimized by a sequential rotation method. In simulations, it is shown that the proposed scheme requires less transmit power than various benchmark methods under the constraint of each user's quality of service.**

***Index Terms*—Coordinated multi-point, group-level successive interference cancellation, non-orthogonal multiple access, reconfigurable intelligent surface.**

I. Introduction

Since 2019, the deployment of the fifth generation (5G) wireless network has been gradually rolled out worldwide. In regard to the development law of "using one generation, developing one generation", the investigation on the sixth generation (6G) communication system has drawn substantial attention from both academic and industrial communities. Several well-known research institutions have issued 6G white papers, wherein reconfigurable intelligent surface (RIS) has been recognized as a promising technique for 6G [1]. Specifically, RIS is a passive device without the capacity of signal processing, which can be easily deployed on building facades, roadside billboards, and so on. By manipulating the phases of incident signals, RIS has the ability of creating favourable radio propagation environment in an energy-efficient and cost-effective manner [2]. It has been reported that both system throughput and power efficiency can be significantly enhanced by the deployment of RISs [3], [4].

In the meanwhile, as the number of mobile terminals is experiencing an exponentially increased rate per year, massive connectivity is forecast to be an essential application for future communications [5]. Due to the limit number of resource blocks (RBs), it is impossible to satisfy the massive access requirements by employing conventional orthogonal multiple access (OMA) techniques. Different from OMA, non-orthogonal multiple access (NOMA) is capable of supporting much more connections than the number of RBs by leveraging superposition coding and successive interference cancellation (SIC) [6]. It was reported by [7], [8] that an improved user fairness and spectrum efficiency can be achieved in NOMA systems by leveraging heterogeneous channel conditions.

As mentioned above, both RIS and NOMA are perceived as cutting-edge techniques for beyond-5G (B5G)/6G communication systems. The integration of RIS into NOMA system induces enormous benefits [9], such as reshaping the difference of NOMA users' channel gains, improving user fairness, and enhancing wireless coverage. It was demonstrated that RIS-empowered NOMA system is powerful to reduce the transmit power consumption or boost the achievable rates compared with the pure NOMA system [10], [11]. After that, enormous attentions have been shifting towards the study of RIS-enabled NOMA systems. Compared with conventional NOMA systems, the RIS's phase shifts should be jointly optimized with the transceivers for RIS-enabled NOMA systems. Therefore, it is nontrivial to devise efficient transmission policies for RIS-assisted NOMA networks. In what follows, we will review the state-of-the-art research on RIS-empowered NOMA networks comprehensively.

A. Related Works

With regard to RIS-assisted NOMA systems, the focuses of the existing works are put on the performance analysis and the design of transmission policy [12], [13]. To highlight the contribution of this paper, we will review the studies on transmission strategy in this subsection.

RIS-Aided Single-Antenna NOMA: In terms of transmission design for RIS-enabled single-antenna NOMA systems, the design methods in existing works can be classified into

This work was supported in part by the National Natural Science Foundation of China under Grants 61771257 and 62171237, in part by the China Postdoctoral Science Foundation under Grants BX20180143 and 2019M660126, and in part by Jiangsu Postdoctoral Science Foundation under Grant 16KJB510035. (*Corresponding Author: Chen Liu*)

H. Wang and R. Song are with the School of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China (e-mail: wanghong@njupt.edu.cn, songrf@njupt.edu.cn).

C. Liu is with the School of Electronic and Optical Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210023, China (e-mail: liuch@njupt.edu.cn).

Z. Shi is with the School of Intelligent Systems Science and Engineering, Jinan University, Zhuhai 519070, China (e-mail: zhengshi@jnu.edu.cn).

Y. Fu is with the School of Science and Technology, Hong Kong Metropolitan University, Hong Kong, China (e-mail: yfu@hkmu.edu.hk).

two main categories, i.e., (i) alternating optimization of phase shifts and power allocation coefficients, and (ii) transforming the joint optimization of phase shifts and power allocation to a pure phase shift optimization problem. Following the first category, the power allocation and the phase shifts were devised in an alternative manner for maximizing system throughput of downlink RIS-aided NOMA systems [14]. In [15], the network sum-rate was maximized by jointly optimizing the power allocation coefficients, the user clustering policy, and the phase shifts for two-cell RIS-aided NOMA downlinks with CoMP transmissions. On the other hand, the sum rate maximization problem was formulated under the constraint of each user's power budget in [16]. To address the joint optimization of power allocation and phase shifts, the original problem was converted into a phase shift determination problem by using a series of mathematical manipulations. However, the transmission strategies developed in [14]–[16] are applicable to RIS-aided single-antenna NOMA systems. In modern communication systems, multiple antennas are deployed at the base station (BS) in general. Unfortunately, the schemes mentioned above are not proper to multi-antenna RIS-NOMA scenarios.

RIS-Aided Multi-Antenna NOMA: In RIS-assisted multi-antenna NOMA systems, the design of transmission schemes is more complicated than that in single-antenna scenario. Even if the phase shifts at the RIS are given, the equivalent channels with the form of vectors/matrices cannot be sorted like that in single-antenna NOMA systems. Therefore, in the case of multi-antenna NOMA situation, a deep integration of NOMA with space division multiple access (SDMA) is necessary. In the existing literature, there are two design frameworks in RIS-aided multi-antenna NOMA systems, i.e., the user-specific and the cluster-specific frameworks. In the following, we will review the prior works from these two perspectives.

In the user-specific framework, each user is served by a unique beam, which is also referred to as beamformer-based strategy in [17]. For two-user RIS-aided multiple-input-single-output (MISO) NOMA downlinks, energy efficiency was enhanced by optimizing passive and active beamformers alternatively in [18]. Particularly, the active beamforming matrix at the BS and the passive beamforming matrix at the RIS were developed by employing successive convex approximation (SCA) and semidefinite relaxing (SDR), respectively. Furthermore, for RIS-assisted MISO-NOMA with multiple users, the beamforming matrix and the phase shifts were collaboratively devised to improve the minimum signal-to-interference-plus-noise-ratio (SINR) by taking each user's quality of service (QoS) into consideration [19]. However, in the user-specific framework, the users decoded later should detect all the users with higher decoding priorities, which will cause severe error propagations in practice.

Regarding the cluster-specific framework, all users are divided into several clusters and the users in each cluster are organized on the same beam by using the NOMA mode, whilst the inter-cluster interference is suppressed by active beamforming at the BS. In [20], a multi-cluster transmission scheme was proposed to reduce the total power consumption for RIS-aided MISO-NOMA systems, in which a local optimal solution was achieved by combing the alternating direction method of multipliers (ADMM) with the second-order cone programming (SOCP). In [21], a multi-cluster downlink NOMA scheme was devised to minimize the transmit power consumption, in which the users are separated into multiple clusters and each cluster is composed of a near user and a far user. In order to enhance the performance of the users with weak channels, the RISs are deployed at the neighbouring of far users.

However, the design frameworks presented in [17]–[21] focused on downlink RIS-aided multi-antenna NOMA system. Because of the different mechanisms on power allocation and decoding order, the downlink design schemes can not be extended to the uplink case explicitly. To date, the investigations on uplink transmission strategies in RIS-assisted multi-antenna NOMA systems are far fewer than those in downlink scenarios. Following the cluster-specific framework, an optimized passive beamforming scheme with NOMA and time allocation was proposed for uplink information transmission of RIS-aided wireless power communication networks in [22]. It was shown that network throughput can be enhanced considerably by using the optimized phase shifts and time allocation in [22].

RIS-Aided Multi-Cell NOMA: The literature mentioned above focused on the single-cell scenarios. In the real-world systems, the cell-edge user's performance is affected by the serious intercell interference due to frequency reuse. In order to deal with this tricky issue, coordinated multi-point (CoMP) technique has been introduced to improve the performance of cell-edge users from 3GPP Release-11, in which joint detection and joint transmission are two well-known techniques respectively applied in the uplink and downlink. In [23], by using tools from stochastic geometry, the average achievable data rate was derived for downlink heterogeneous cloud radio access networks with NOMA and CoMP transmission. In [24], both the outage probabilities and the ergodic rates were derived for uplink CoMP-NOMA networks with the aid of stochastic geometry. However, in [23], [24], only one CoMP user was served by each BS equipped with a single antenna. In [25], a sum-rate maximization problem was investigated for downlink RIS-aided CoMP networks with device-to-device communications, which was further solved by leveraging the SCA and the penalty method. In [26], a resource allocation problem was investigated for downlink heterogeneous MISO-NOMA networks with CoMP, which was further addressed by the method based on matching algorithms and the SCA.

After the emergence of RIS, it was integrated into multi-cell CoMP system to boost cell-edge user's transmission rate in [27]. It was shown that the sum rate is double of the no-RIS case when the number of reflecting units is more than 80. However, the conventional OMA mode was considered in [27], which is fundamentally different from the case of NOMA systems. In multi-cell scenarios, the existing transmission designs are still limited to RIS-assisted single-antenna NOMA systems. More specifically, in [28], a joint power control and phase shift optimization problem was converted into a pure phase shift determination problem for RIS-aided multi-cell NOMA uplinks, which was further addressed by leveraging the approaches of angle minimization and Schmidt's orthogonalization. Besides, a RIS was applied to improve the SINRs of

This article has been accepted for publication in IEEE Transactions on Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2022.3177773

cell-edge users in [29], in which the users' ergodic capacities were derived in closed-forms.

B. Our Contributions

It has been demonstrated that the CoMP-NOMA system is superior to the conventional CoMP system in terms of power efficiency for a two-cell network [30], in which CoMP has been advocated from 3GPP Release 11. Besides, since RIS has the benefits of signal enhancement and neutralization, it is integrated into two-cell CoMP-NOMA systems to further boost system performance in this work. Currently, the transmission design for RIS-aided multi-antenna NOMA is still in its infancy. To the best of our knowledge, the research on transmission scheme for uplink RIS-assisted multi-antenna CoMP-NOMA networks has never been reported, which however, is an important application for future communication networks. This constitutes the motivation of our work. Specifically, in this paper, the transmission strategy is studied for CoMP-NOMA networks with multiple user groups, in which both RIS and CoMP are leveraged to boost the cell-edge users' signal strengthes. In the authors' prior work [31], a single cell with distributed RIS deployment is investigated, whereas a two-cell CoMP network with centralized RIS deployment is considered in this paper. Since the perfect SIC is assumed in [31], the power control schemes are identical for all user groups, which are different from the case of imperfect SIC in this work. Besides, the design of phase shifts in [31] considers the constraints for a local RIS group only. Thus, the power allocation and phase shift algorithms developed in [31] can not be applied to this work. To summarize, the main contributions of this work can be listed as follows:

- A novel design framework with group-level SIC (GSIC) is conceived for RIS-aided CoMP-NOMA networks with multiple user groups, in which GSIC is proposed to separate the user messages in different groups and the transceivers is used to suppress the aggregated interferences, i.e., the intra-group interference, the residual interference caused by imperfect SIC, and the interference of NOMA users decoded later.[1] Different from the user-level SIC in the existing frameworks, GSIC is leveraged to eliminate the decoded group's interference for the first time.[2] It is worth mentioning that GSIC is more appropriate for multi-group NOMA networks because of homogeneous channel conditions for the users in the same group.
- With this new design framework, a power minimization problem is formulated by jointly considering the power control scheme at the users, the equalizers at the BS, and the phase shifts at the RIS. Besides, the impact of imperfect SIC is also included in the problem formulation, in which additional interference caused by the previously decoded users is considered. Since the optimization variables are highly coupled, it is nontrivial to solve the original non-convex problem explicitly. To render the formulated problem tractable, an alternating optimization method is developed to determine the transmit power of each group and the phase shifts of the RIS.
- More specifically, a group-level parallel iteration method is proposed to devise each user's transmit power. The designed solution is proved to converge to a unique stationary point for an arbitrary initialization value. Besides, a sequential phase rotation method is developed to optimize each phase shift at the RIS. Simulation results demonstrate that our proposed scheme outperforms the existing methods in terms of transmit power consumption.

The rest of this paper is organized as follows. In Section II, the system model and detection procedures are described. The optimization problem and its solution are respectively formulated and proposed in Section III. Section IV elaborates the simulation results and their analyses. The conclusion is given in Section V.

Notations: Unless otherwise stated, throughout this paper, lowercase and uppercase boldface letters denote vectors and matrices, respectively. $\text{diag}\{z_1, \cdots, z_m, \cdots, z_M\}$ denotes a diagonal matrix with the m-th diagonal element of z_m. $[\mathbf{z}]_m$ denotes the m-th element of the vector $\mathbf{z}$. $\mathbf{Z}^T$, $\mathbf{Z}^{-1}$, $\mathbf{Z}^*$, and $\mathbf{Z}^H$ denote the transpose, inversion, complex conjugate, and Hermitian transpose of the matrix $\mathbf{Z}$, respectively. $\mathbf{z}_1 \succeq \mathbf{z}_2$ represents that $[\mathbf{z}_1]_m \geq [\mathbf{z}_2]_m$ for $\forall m$. $\mathbf{I}_M$ represents an identity matrix with M rows and M columns. $\mathbf{0}_{M_1 \times M_2}$ denotes a zero matrix with M_1 rows and M_2 columns. mod stands for the modulo operation. $\jmath = \sqrt{-1}$.

II. System Model

As illustrated in Fig. 1, the uplink transmission of two-cell RIS-enabled CoMP-NOMA networks is considered, in which each BS serves a cell-center user group and a cell-edge user group.[3] The coverage radius of each BS is R_B. A RIS with

[1] In the proposed GSIC framework, the interference can be divided into three types, i.e., the intra-group interference, the residual interference, and the interference of NOMA users decoded later. Specifically, the intra-group interference is caused by the users in the same group due to parallel demodulation, the residual interference comes from the users of the decoded groups because of imperfect SIC, and the interference of NOMA users decoded later stems from the user groups that have not been decoded.

[2] For the NOMA with user-level SIC, the users are divided into multiple clusters, which are served by different space beams, time slots, or frequency bands. In each cluster, the user-level SIC is used to eliminate the user's signal that has been decoded. However, the GSIC is leveraged to eliminate each group's signals, which are decoded and removed in parallel. Besides, for the NOMA with GSIC, the equalizers, power control, and phase shifts are jointly optimized to suppress intra-group and inter-group interference for the users in the same group, whereas the users in the different groups are served by following the NOMA principle and the GSIC is invoked to subtract the decoded group's messages.

[3] It is worth mentioning that, in multi-cell scenarios, any two adjacent cells can form a CoMP partner similar to the system architecture in Fig. 1. When the different NOMA clusters in the same cell are scheduled on orthogonal subchannels or time slots, the design schemes proposed in this paper can be extended to RIS-aided multi-cell CoMP-NOMA scenarios straightforwardly. A feasible user clustering scheme is provided as follows. The cell-edge users associated with each RIS are assigned to the same cell-edge group. For the cell-center users in each cell, they could be allocated to multiple NOMA clusters randomly under the constraint that the number of users in each cell-center group is less than that of the antennas at the BS. In the case of the cell-edge users associated with multiple BSs, the cell-edge users' signals at each associated BS will be forwarded to the central unit for joint detection. As a result, the SINR expression of the cell-edge user given in (15) will be formulated by considering the combination of the desired signals and the cell-center users' interference stemming from multiple BSs.

This article has been accepted for publication in IEEE Transactions on Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2022.3177773

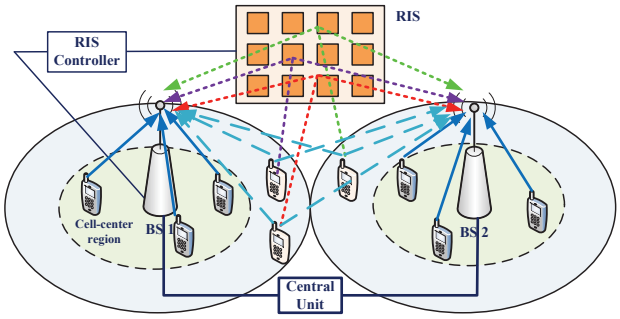


Fig. 1. An illustration of uplink transmissions of RIS-assisted multi-group NOMA with CoMP.

N_S reflecting elements is deployed at the edge of the two cells to manipulate the phases of the cell-edge users' signals. The coverage radii of the cell-center region and the RIS are R_C and R_S, respectively. The number of antennas equipped at the BS is N_A, while each user is equipped with a single antenna. The cell-center groups are served by their associated BS independently, whilst the cell-edge group is associated with the two BSs by using CoMP reception. Specifically, after the cell-center group's signals are recovered and subtracted at each BS, the users' signals of the cell-edge group are forwarded to the central unit (CU) for joint detection. It is worth mentioning that the CU can be an independent entity or any of the two BSs with powerful signal processing capability. To do so, a high-capacity backhaul link (fiber or wireless) should be supported between the BSs and the CU, such as S1 interface in current communication systems. Besides, it is assumed that the number of active users in the cell-center group of BS b and the cell-edge group are $K_{C,b}$ and K_E, respectively. In order to decode the data streams in each group in parallel, it is required that $K_{C,b} \leq N_A$ and $K_E \leq 2N_A$.[4] In addition, we consider the scenario that the number of active users scheduled by each BS is more than that of the receiving antennas at the BS, i.e.,

$$K_{C,b} + K_E > N_A, \ b \in \{1, 2\}. \tag{1}$$

In such circumstance, the user signals can not be recovered by the conventional multi-user receivers of OMA systems.

Regarding the combination of NOMA, RIS, and CoMP techniques, it does not require large changes in the circuits of current communication systems. For NOMA, a SIC receiver should be equipped at the BS. With the development of circuits, it is not difficult to satisfy such requirement. For RIS, it requires a RIS controller between the BS and the RIS to adjust the phase shifts. Besides, in order to jointly design the transmit power and phase shifts, the channel estimation of reflection links is needed, which will increase the system overhead. Fortunately, a quantity of high-quality channel estimation methods have been developed in the existing works. The challenges of this combined RIS-aided CoMP-NOMA system are two-fold. Firstly, a high-speed backhaul link should be supported between the BSs and the CU. In practice, the received signals can be quantized and compressed in order to reduce the feedback overhead [32]. Secondly, it is difficult to acquire the perfect channel state information (CSI) in the scheme design. Thus, the power consumption obtained in this paper can serve as a lower bound for the cases with imperfect signal feedback or CSI.

[4]If these conditions are not satisfied, the extra users can be scheduled on the orthogonal time slots or frequencies via OMA methods. Therefore, the method developed in this work is also applicable to the case of $K_{C,b} > N_A$ or $K_E > 2N_A$.

A. Channel Model

The direct link between BS b and the i-th user in the cell-center or the cell-edge groups are modelled as Raleigh distribution. Besides, the Kronecker model is used to characterize spatial correlation among the receiving antennas at the BS or the reflectors at the RIS [33]. The direct link between BS b and the i-th user in the cell-center and the cell-edge groups are given as

$$\mathbf{d}_{C,b,i} = L_{C,b,i}^{-\frac{\beta_{B,U}}{2}} \boldsymbol{\Upsilon}_{B,R}^{\frac{1}{2}} \mathbf{f}_{C,b,i}, \ 1 \leq i \leq K_{C,b}, \tag{2}$$

$$\mathbf{d}_{E,b,i} = L_{E,b,i}^{-\frac{\beta_{B,U}}{2}} \boldsymbol{\Upsilon}_{B,R}^{\frac{1}{2}} \mathbf{f}_{E,b,i}, \ 1 \leq i \leq K_E, \tag{3}$$

respectively, where $L_{C,b,i}$ and $L_{E,b,i}$ denote the propagation distances from the user i in the cell-center group to the BS b and from the user i in the cell-edge group to the BS b, respectively, $\beta_{B,U}$ stands for the path loss exponent for the direct channels between the BSs and their served users, each element in the small-scale fading vector $\mathbf{f}_{C,b,i}$ or $\mathbf{f}_{E,b,i}$ follows a complex Gaussian distribution with zero mean and unit variance, i.e., $\mathcal{CN}(0,1)$, and $\boldsymbol{\Upsilon}_{B,R}$ represents the receiving correlation matrix at the BS.

In order to improve the channel gains, the RIS can be deployed at a carefully selected location, where there exists a line-of-sight (LoS) path between the RIS and the users (the BSs). It is worth mentioning that the location of the RIS only affects channel modeling in this subsection, which will not influence the design methods of transceivers and phase shifts developed in the following. Accordingly, the channel between the RIS and the i-th cell-edge user in its coverage is characterized by Rician distribution, which is given by

$$\mathbf{g}_{S,i} = L_{S,i}^{-\frac{\beta_{S,U}}{2}} \left(\sqrt{\frac{\kappa_{S,U}}{\kappa_{S,U}+1}} \bar{\mathbf{f}}_{S,i} + \sqrt{\frac{1}{\kappa_{S,U}+1}} \boldsymbol{\Upsilon}_{S,R}^{\frac{1}{2}} \mathbf{f}_{S,i} \right), \tag{4}$$

where $L_{S,i}$ denotes the propagation distance from the i-th cell-edge user to the RIS, $\kappa_{S,U}$ is the Rician-K factor of the channels between the RIS and its associated users, $\beta_{S,U}$ represents the path loss exponent for the links from the cell-edge group to the RIS, $\bar{\mathbf{f}}_{S,i}$ is the LoS component from the i-th cell-edge user to the RIS, each element in the vector $\mathbf{f}_{S,i}$ follows $\mathcal{CN}(0,1)$, and $\boldsymbol{\Upsilon}_{S,R}$ stands for the receiving correlation matrix at the RIS.

Similarly, the channel from the RIS to the BS b is characterized by

$$\mathbf{G}_{b,S} = L_{b,S}^{-\frac{\beta_{B,S}}{2}} \left(\sqrt{\frac{\kappa_{B,S}}{\kappa_{B,S}+1}} \bar{\mathbf{F}}_{b,S} + \sqrt{\frac{1}{\kappa_{B,S}+1}} \boldsymbol{\Upsilon}_{B,R}^{\frac{1}{2}} \mathbf{F}_{b,S} \boldsymbol{\Upsilon}_{S,T}^{\frac{1}{2}} \right), \tag{5}$$

where $L_{b,S}$ denotes the propagation distance between the RIS and the BS b, $\kappa_{S,U}$ is the Rician-K factor of the channels between the RIS and the BSs, $\beta_{B,S}$ stands for the path loss exponent between the RIS and the BSs, $\bar{\mathbf{F}}_{b,S}$ represents the

LoS component for the channels between the RIS and the BSs, each element in $\mathbf{F}_{b,S}$ follows $\mathcal{CN}(0,1)$, and $\boldsymbol{\Upsilon}_{S,T}$ denotes the reflecting correlation matrix at the RIS.

B. User Detection in the Cell-Center Group

Due to severe obstacle blockage and substantial path attenuation, the signals reflected by two or more times and the interference stemming from the cell-center group to the unconnected BS are neglected [34]. Besides, due to product-distance path loss of cascaded channels, the cell-center users' signals reflected by the RIS with large propagation distances are much weaker than those of the direct links. Thus, it is reasonable to neglect the cell-center group's signals reflected by the RIS [34], [35]. Accordingly, the received signal at BS b is given by

$$\mathbf{z}_b = \sum_{i=1}^{K_{C,b}} \mathbf{d}_{C,b,i}\sqrt{P_{b,i}}m_{b,i} + \sum_{i=1}^{K_E} \mathbf{c}_{E,b,i}\sqrt{P_{E,i}}m_{E,i} + \mathbf{n}_b, \quad b \in \{1,2\}, \tag{6}$$

where $P_{b,i}$ and $P_{E,i}$ are the transmit power of user i in the cell-center and the cell-edge groups, respectively, $m_{b,i}$ and $m_{E,i}$ represent the transmit messages with normalized power of the user i in the cell-center group and the cell-edge group, respectively, $\mathbf{n}_b$ is the additive noise vector at BS b, each element of which follows $\mathcal{CN}(0,\sigma^2)$, and $\mathbf{c}_{E,b,i}$ stands for the cascaded channel between user i of the cell-edge group and the BS, which is further expressed as

$$\mathbf{c}_{E,b,i} = \mathbf{d}_{E,b,i} + \mathbf{G}_{b,S}\boldsymbol{\Phi}\mathbf{g}_{S,i}, \tag{7}$$

where $\boldsymbol{\Phi}$ is the phase shift matrix of the RIS, which can be expressed as $\boldsymbol{\Phi} = \mathrm{diag}\{e^{\jmath\omega_1}, \cdots, e^{\jmath\omega_{N_S}}\}$ with ω_n denoting the phase shift of the n-th reflecting unit. From the perspective of hardware limitation, the discrete phase shift model is considered in this paper. The set of discrete phase shifts is generated by applying uniform quantization methods [36]. Assuming the number of resolution bits is X, the set of phase shifts is expressed as

$$\mathcal{X} = \left\{0/2^X, 2\pi/2^X, \cdots, (2^X-1)2\pi/2^X\right\}, \tag{8}$$

where $\omega_n \in \mathcal{X}$ for $1 \le n \le N_S$.

Due to the small path attenuation, the signals of the cell-center groups are decoded firstly at each BS.[5] When decoding the messages of the cell-center group, the cell-edge group's signals are treated as the interference. Besides, the users in the cell-center group are detected in parallel. Let $\mathbf{t}_{b,i}$ denote the equalizer for decoding the signal of user i in the cell-center group. Then, the decoded data of user i in the cell-center group is expressed as

$$\begin{aligned}\hat{m}_{b,i} =& \mathbf{t}_{b,i}^H \sum_{i'=1}^{K_{C,b}} \mathbf{d}_{C,b,i'}\sqrt{P_{b,i'}}m_{b,i'} \\ &+ \mathbf{t}_{b,i}^H \sum_{i'=1}^{K_E} \mathbf{c}_{E,b,i'}\sqrt{P_{E,i'}}m_{E,i'} + \mathbf{t}_{b,i}^H\mathbf{n}_b.\end{aligned} \tag{9}$$

In the proposed framework, the users' signals in the same group are decoded in parallel by using transceiver design, while the GSIC method is utilized to eliminate the decoded group's interference. Therefore, in the decoding of a cell-center user's signal, there still exists intra-group interference caused by the other cell-center users due to parallel demodulation. As a result, the achievable SINR of user i in the cell-center group is expressed as

$$\eta_{b,i} = \frac{|\mathbf{t}_{b,i}^H\mathbf{d}_{C,b,i}|^2 P_{b,i}}{\sum\limits_{i'=1,i'\neq i}^{K_{C,b}} |\mathbf{t}_{b,i}^H\mathbf{d}_{C,b,i'}|^2 P_{b,i'} + \sum\limits_{i'=1}^{K_E} |\mathbf{t}_{b,i}^H\mathbf{c}_{E,b,i'}|^2 P_{E,i'} + \sigma^2\mathbf{t}_{b,i}^H\mathbf{t}_{b,i}}. \tag{10}$$

C. User Detection in the Cell-Edge Group

When the user detection of the cell-center group is completed, the cell-center users' signals are recovered and subtracted from each BS's total received signals. Due to the error propagation caused by imperfect SIC, there exists the residual interference stemming from the cell-center users in practice. After the operation of GSIC, the remaining signal at BS b is expressed by

$$\begin{aligned}\mathbf{z}_{E,b} =& \sum_{i=1}^{K_E} \mathbf{c}_{E,b,i}\sqrt{P_{E,i}}m_{E,i} + \sum_{i=1}^{K_{C,b}} \sqrt{\varepsilon_{b,i}}\mathbf{d}_{C,b,i}\sqrt{P_{b,i}}m_{b,i} \\ &+ \mathbf{n}_b,\end{aligned} \tag{11}$$

where $\varepsilon_{b,i}$ denotes the residual interference factor caused by the cell-center user i served by BS b. Since the messages of the cell-edge group will be decoded at the CU jointly, each BS transfers the remaining signal $\mathbf{z}_{E,b}$ to the CU without any preprocessing.

For facilitating joint detection at the CU, the residual signals coming from the two BSs are written in a compact form, which is given by

$$\begin{aligned}\mathbf{z}_E =& \begin{bmatrix} \mathbf{z}_{E,1} \\ \mathbf{z}_{E,2} \end{bmatrix} \\ =& \sum_{i=1}^{K_E} \mathbf{c}_{E,i}\sqrt{P_{E,i}}m_{E,i} + \sum_{b=1}^{2}\sum_{i=1}^{K_{C,b}} \sqrt{\varepsilon_{b,i}}\tilde{\mathbf{d}}_{C,b,i}\sqrt{P_{b,i}}m_{b,i} \\ &+ \mathbf{n}_E,\end{aligned} \tag{12}$$

where

$$\mathbf{c}_{E,i} = \begin{bmatrix} \mathbf{d}_{E,1,i} \\ \mathbf{d}_{E,2,i} \end{bmatrix} + \begin{bmatrix} \mathbf{G}_{1,S} \\ \mathbf{G}_{2,S} \end{bmatrix} \boldsymbol{\Phi}\mathbf{g}_{S,i}, \tag{13}$$

$$\tilde{\mathbf{d}}_{C,1,i} = \begin{bmatrix} \mathbf{d}_{C,1,i} \\ \mathbf{0}_{N_A\times 1} \end{bmatrix}, \tilde{\mathbf{d}}_{C,2,i} = \begin{bmatrix} \mathbf{0}_{N_A\times 1} \\ \mathbf{d}_{C,2,i} \end{bmatrix}, \mathbf{n}_E = \begin{bmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \end{bmatrix}. \tag{14}$$

[5]Since the signals of the cell-center group are decoded at each BS, the CoMP reception is implemented after the GSIC operation. Thus, CoMP does not affect the NOMA detection at each BS. In the presence of multiple antennas, the channels are in the form of vectors/matrices. Thus, it is hard to sort the channels like the case with a single antenna [7]. Currently, the optimal detection order in multi-antenna NOMA system is still an open problem. In this paper, the detection order is determined by the average path loss for each cell. In fact, in order to achieve the required data rates, the introduction of RIS will reduce the power consumption of cell-edge users for uplink NOMA systems, which will in turn reduce the inter-NOMA user interference to the cell-center users. Therefore, the introduction of RIS will not affect the feasibility of signal decoding but reduce the required transmit power.

At the CU, the equalizer of the user i in the cell-edge group is denoted by $\mathbf{t}_{E,i}$. Similar to the decoding procedures of the cell-center group, the achievable SINR of user i in the cell-edge group is expressed as

$$\eta_{E,i} = \frac{|\mathbf{t}_{E,i}^H \mathbf{c}_{E,i}|^2 P_{E,i}}{\sum\limits_{i'=1,i'\neq i}^{K_E} |\mathbf{t}_{E,i}^H \mathbf{c}_{E,i'}|^2 P_{E,i'} + \sum\limits_{b=1}^{2}\sum\limits_{i'=1}^{K_{C,b}} |\mathbf{t}_{E,i}^H \tilde{\mathbf{d}}_{C,b,i'}|^2 \varepsilon_{b,i'} P_{b,i'} + \sigma^2 \mathbf{t}_{E,i}^H \mathbf{t}_{E,i}}. \tag{15}$$

III. Problem Formulation and Its Solution

In order to satisfy the urgent global carbon reduction consensus, green communication is one of the main themes for B5G/6G communication networks. In this section, we will investigate the power minimization problem for RIS-empowered CoMP-NOMA networks subject to each user's QoS. The optimization problem is formulated as follows:

$$\text{(P1)} \quad \min_{\{P_{b',i}\},\{\mathbf{t}_{b',i}\},\mathbf{\Phi}} : \sum_{b=1}^{2}\sum_{i=1}^{K_{C,b}} P_{b,i} + \sum_{i=1}^{K_E} P_{E,i}, \tag{16}$$

$$\text{s.t. } \log_2(1+\eta_{b',i}) \geq \gamma_{b',i}^{(\text{th})}, \ b' \in \{1,2,E\}, \tag{17}$$

$$P_{b',i} \leq P_{\max}, \ b' \in \{1,2,E\}, \tag{18}$$

$$\omega_n \in \mathcal{X}, \ 1 \leq n \leq N_S, \tag{19}$$

where $\gamma_{b',i}^{(\text{th})}$ represents the required transmit rate of user i in the cell-center or the cell-edge groups, $P_{\max}$ denotes the allowable transmission power of each user. In the problem (P1), the 1st constraint is used to restrict each user's transmission rate larger than a predefined threshold, the 2nd constraint ensures the transmit power is less than the maximum allowable power, and the 3rd constraint requires that each phase shift is selected from a predefined set $\mathcal{X}$.

A. The Equalizer Design and Problem Transformation

In order to realize each user's SINR maximization, the minimum mean square error (MMSE) equalizer is recognized as the best linear receiver. The optimal receiver for the i-th cell-center user associated with BS b is written as a function of transmit powers, which is given by

$$\mathbf{t}_{b,i} = \left(\mathcal{F}_{C,b}(\mathbf{p}_b, \mathbf{p}_E)\right)^{-1} \mathbf{d}_{C,b,i} \sqrt{P_{b,i}}, \tag{20}$$

where $\mathbf{p}_b = \{P_{b,i'}\}$, $\mathbf{p}_E = \{P_{E,i'}\}$, and $\mathcal{F}_{C,b}(\mathbf{p}_b, \mathbf{p}_E)$ is defined as follows:

$$\mathcal{F}_{C,b}(\mathbf{p}_b, \mathbf{p}_E) \triangleq \sum_{i'=1}^{K_{C,b}} \mathbf{d}_{C,b,i'} \mathbf{d}_{C,b,i'}^H P_{b,i'} + \sum_{i'=1}^{K_E} \mathbf{c}_{E,b,i'} \mathbf{c}_{E,b,i'}^H P_{E,i'} + \sigma^2 \mathbf{I}_{N_A}. \tag{21}$$

By substituting the optimal receiver of (20) into (10), the SINR of user i in the cell-center group can be reexpressed as

$$\eta_{b,i} = \mathbf{d}_{C,b,i}^H \left(\mathcal{F}_{C,b}(\mathbf{p}_b, \mathbf{p}_E) - \mathbf{d}_{C,b,i} \mathbf{d}_{C,b,i}^H P_{b,i}\right)^{-1} \mathbf{d}_{C,b,i} P_{b,i}. \tag{22}$$

Analogically, the optimal receiver for the i-th user of the cell-edge group is written as

$$\mathbf{t}_{E,i} = \left(\mathcal{F}_E(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_E)\right)^{-1} \mathbf{c}_{E,i} \sqrt{P_{E,i}}, \tag{23}$$

with

$$\mathcal{F}_E(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_E) \triangleq \sum_{b=1}^{2}\sum_{i'=1}^{K_{C,b}} \tilde{\mathbf{d}}_{C,b,i'} \tilde{\mathbf{d}}_{C,b,i'}^H \varepsilon_{b,i'} P_{b,i'} + \sum_{i'=1}^{K_E} \mathbf{c}_{E,i'} \mathbf{c}_{E,i'}^H P_{E,i'} + \sigma^2 \mathbf{I}_{2N_A}. \tag{24}$$

By substituting the optimal receiver of (23) into (15), the SINR of user i in the cell-edge group is further derived by

$$\eta_{E,i} = \mathbf{c}_{E,i}^H \left(\mathcal{F}_E(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_E) - \mathbf{c}_{E,i} \mathbf{c}_{E,i}^H P_{E,i}\right)^{-1} \mathbf{c}_{E,i} P_{E,i}. \tag{25}$$

By using the simplified SINR expressions, the optimization problem (P1) is transformed to

$$\text{(P2)} \quad \min_{\{P_{b',i}\},\mathbf{\Phi}} : \sum_{b=1}^{2}\sum_{i=1}^{K_{C,b}} P_{b,i} + \sum_{i=1}^{K_E} P_{E,i} \tag{26}$$

$$\text{s.t. } \log_2(1+\eta_{b',i}) \geq \gamma_{b',i}^{(\text{th})}, \ b' \in \{1,2,E\}, \tag{27}$$

$$P_{b',i} \leq P_{\max}, \ b' \in \{1,2,E\}, \tag{28}$$

$$\omega_n \in \mathcal{X}, \ 1 \leq n \leq N_S. \tag{29}$$

Since the feasible set is non-convex, the optimization problem (P2) is non-convex accordingly. Besides, the optimization variables $\{P_{b',i}\}$ and $\{\omega_n\}$ are highly coupled in the transformed problem. Thus, it is quite challenging to address the problem (P2) directly. To render the problem tractable, an alternating scheme is proposed to find a high-quality solution, i.e., the transmit power and the phase shifts are optimized alternatively. Specifically, the implementation procedure is given by: $\cdots \rightarrow \{\{P_{b',i}^{(t-1)}\}, \{\omega_n^{(t-1)}\}\} \rightarrow \{\{P_{b',i}^{(t)}\}, \{\omega_n^{(t-1)}\}\} \rightarrow \{\{P_{b',i}^{(t)}\}, \{\omega_n^{(t)}\}\} \rightarrow \cdots$, where t denotes the iteration index. When the objective function value is convergent, the alternating method is terminated. In the following two subsections, we will focus on the design of transmit power coefficients and phase shifts in the t-th iteration.

B. Transmit Power Optimization

In this subsection, the optimal power control scheme is obtained based on the optimized values in the last outer iteration, i.e., $\{\mathbf{\Phi}^{(t-1)}, P_{b,i}^{(t-1)}\}$ are given. Accordingly, the optimization problem (P2) is simplified as

$$\text{(P3)} \quad \min_{\{P_{b',i}\}} : \sum_{b=1}^{2}\sum_{i=1}^{K_{C,b}} P_{b,i} + \sum_{i=1}^{K_E} P_{E,i} \tag{30}$$

$$\text{s.t. } \log_2(1+\eta_{b',i}) \geq \gamma_{b',i}^{(\text{th})}, \ b' \in \{1,2,E\}, \tag{31}$$

$$P_{b',i} \leq P_{\max}, \ b' \in \{1,2,E\}. \tag{32}$$

Proposition 1. *When the optimization problem (P3) obtains the optimal solution, the equalities are achieved in the constraint* (31)*, i.e.,*

$$\eta_{b',i} = 2^{\gamma_{b',i}^{(th)}} - 1, \ b' \in \{1,2,E\}. \tag{33}$$

Proof: This proposition is proved by invoking the contradiction. First, we prove that the equality holds for the cell-center groups. Assume $\hat{\mathbf{p}}_b = \{\hat{P}_{b,i}\}$ to be the optimal transmit powers of the cell-center users in problem (P3), and when the problem gets the optimal solution, we have

$$\eta_{b,i} > 2^{\gamma_{b,i}^{(\text{th})}} - 1. \tag{34}$$

Accordingly, we have

$$\begin{aligned}&\mathbf{d}_{C,b,i}^H \left(\mathcal{F}_{C,b}(\hat{\mathbf{p}}_b, \mathbf{p}_E) - \mathbf{d}_{C,b,i}\mathbf{d}_{C,b,i}^H \hat{P}_{b,i}\right)^{-1} \mathbf{d}_{C,b,i}\hat{P}_{b,i} \\ >&2^{\gamma_{b,i}^{(\text{th})}} - 1,\end{aligned} \tag{35}$$

where $\mathbf{p}_E = \{P_{E,i}\}$. Here, we define a new variable as

$$\begin{aligned}\tilde{P}_{b,i} &= \frac{2^{\gamma_{b,i}^{(\text{th})}} - 1}{\mathbf{d}_{C,b,i}^H \left(\mathcal{F}_{C,b}(\hat{\mathbf{p}}_b, \mathbf{p}_E) - \mathbf{d}_{C,b,i}\mathbf{d}_{C,b,i}^H \hat{P}_{b,i}\right)^{-1} \mathbf{d}_{C,b,i}} \\ &< \hat{P}_{b,i}.\end{aligned} \tag{36}$$

It can be derived that the newly defined variable satisfies the following equation:

$$\begin{aligned}\tilde{\eta}_{b,i} &= \mathbf{d}_{C,b,i}^H \left(\mathcal{F}_{C,b}(\tilde{\mathbf{p}}_b, \mathbf{p}_E) - \mathbf{d}_{C,b,i}\mathbf{d}_{C,b,i}^H \tilde{P}_{b,i}\right)^{-1} \mathbf{d}_{C,b,i}\tilde{P}_{b,i} \\ &= 2^{\gamma_{b,i}^{(\text{th})}} - 1,\end{aligned} \tag{37}$$

where $\tilde{\mathbf{p}}_b \triangleq \{\hat{P}_{b,1}, \cdots, \hat{P}_{b,i-1}, \tilde{P}_{b,i}, \hat{P}_{b,i+1}, \cdots, \hat{P}_{b,K_{C,b}}\}$.

For the transmit power vector $\tilde{\mathbf{p}}_b$, the SINR of the cell-center user j ($j \neq i$) is derived as

$$\begin{aligned}\tilde{\eta}_{b,j} =&\mathbf{d}_{C,b,j}^H \left(\mathcal{F}_{C,b}(\tilde{\mathbf{p}}_b, \mathbf{p}_E) - \mathbf{d}_{C,b,j}\mathbf{d}_{C,b,j}^H \hat{P}_{b,j}\right)^{-1} \mathbf{d}_{C,b,j}\hat{P}_{b,j} \\ =&\mathbf{d}_{C,b,j}^H \left(\mathcal{F}_{C,b,\bar{j}}(\tilde{\mathbf{p}}_b, \mathbf{p}_E) - \mathbf{d}_{C,b,i}\mathbf{d}_{C,b,i}^H \tilde{P}_{b,i}\right. \\ &\left.+\mathbf{d}_{C,b,i}\mathbf{d}_{C,b,i}^H \tilde{P}_{b,i}\right)^{-1} \mathbf{d}_{C,b,j}\hat{P}_{b,j} \\ =&\mathbf{d}_{C,b,j}^H \left(\mathcal{F}_{C,b,\bar{j}}(\tilde{\mathbf{p}}_b, \mathbf{p}_E) - \mathbf{d}_{C,b,i}\mathbf{d}_{C,b,i}^H \tilde{P}_{b,i}\right)^{-1} \mathbf{d}_{C,b,j}\hat{P}_{b,j}, \\ &- \frac{\hat{P}_{b,j}\left[\mathbf{d}_{C,b,j}^H \left(\mathcal{F}_{C,b,\bar{j}}(\tilde{\mathbf{p}}_b, \mathbf{p}_E) - \mathbf{d}_{C,b,i}\mathbf{d}_{C,b,i}^H \tilde{P}_{b,i}\right)^{-1} \mathbf{d}_{C,b,i}\right]^2}{\frac{1}{\tilde{P}_{b,i}} + \mathbf{d}_{C,b,i}^H \left(\mathcal{F}_{C,b,\bar{j}}(\tilde{\mathbf{p}}_b, \mathbf{p}_E) - \mathbf{d}_{C,b,i}\mathbf{d}_{C,b,i}^H \tilde{P}_{b,i}\right)^{-1} \mathbf{d}_{C,b,i}},\end{aligned} \tag{38}$$

with

$$\mathcal{F}_{C,b,\bar{j}}(\tilde{\mathbf{p}}_b, \mathbf{p}_E) \triangleq \mathcal{F}_{C,b}(\tilde{\mathbf{p}}_b, \mathbf{p}_E) - \mathbf{d}_{C,b,j}\mathbf{d}_{C,b,j}^H \hat{P}_{b,j},$$

where the last equation in (38) holds by following the Sherman-Morrison formula [37].

Based on the entries at the right hand of the last equation in (38), it is obvious that the SINR $\tilde{\eta}_{b,j}$ is increased monotonously with the decreasing of $\tilde{P}_{b,i}$. Thus, we have

$$\begin{aligned}\tilde{\eta}_{b,j} > \mathbf{d}_{C,b,j}^H &\left(\mathcal{F}_{C,b}(\hat{\mathbf{p}}_b, \mathbf{p}_E) - \mathbf{d}_{C,b,j}\mathbf{d}_{C,b,j}^H \hat{P}_{b,j}\right)^{-1} \mathbf{d}_{C,b,j}\hat{P}_{b,j} \\ &\geq 2^{\gamma_{b,j}^{(\text{th})}} - 1.\end{aligned} \tag{39}$$

Besides, it can be derived that the SINRs of the cell-edge group are non-decreased in the same way. According to (37) and (39), we can get that $\tilde{\mathbf{p}}_b$ is also a feasible solution to optimization problem (P3). Furthermore, by recalling (36), we can get

$$\sum_{i'=1, i' \neq i}^{K_{C,b}} \{\hat{P}_{b,i'}\} + \tilde{P}_{b,i} < \sum_{i'=1}^{K_{C,b}} \{\hat{P}_{b,i'}\}. \tag{40}$$

Thus, we can get that $\hat{\mathbf{p}}_b$ is not the optimal solution of problem (P3), which contradicts with the above assumption that $\hat{\mathbf{p}}_b$ is the optimal solution to the problem. Similarly, the equalities in the constraint of the cell-edge users can be proved accordingly, whose proof is omitted to avoid redundancy. Here, the proof of Proposition 1 is completed. ■

Proposition 2. *If the problem (P3) is feasible, the optimal solution of this problem is equivalent to the problem (P4), which is given as*

$$\textit{(P4)} \quad \min_{\{P_{b',i}\}} : \sum_{b=1}^{2} \sum_{i=1}^{K_{C,b}} P_{b,i} + \sum_{i=1}^{K_E} P_{E,i} \tag{41}$$

$$s.t. \log_2(1 + \eta_{b',i}) = \gamma_{b',i}^{(th)}, \ b' \in \{1, 2, E\}. \tag{42}$$

Proof: We prove this proposition by considering the following two cases.

Case 1: The optimal transmit powers obtained by problem (P4) are less than $P_{\max}$, i.e., $P_{b',i} \leq P_{\max}, \ \forall b', i$. In such case, it is obvious that problem (P3) is feasible and its optimal solution is the same as that of problem (P4).

Case 2: When problem (P4) gets the optimal solution, some transmit powers are larger than $P_{\max}$, e.g., $P_{b',j} > P_{\max}$. In this case, if the rate constraints of other users still hold, there is no way to reduce the transmit power of user j in order to satisfy the constraint $\log_2(1+\eta_{b',j}) = \gamma_{b',j}^{(\text{th})}$. Then, it is easy to obtain that problem (P3) is infeasible. Therefore, Proposition 2 is proved. ■

Thus, based on the above analyses, a parallel iteration method is used to obtain the optimal solution of problem (P4). In the design of each user's transmit power, a group-level power updating scheme is developed. Specifically, the transmit power coefficients of the cell-center group are devised for each cell. After that, the power control of the cell-edge group is implemented based on the designed power coefficients of the cell-center groups. The power control algorithm of the cell-center group is expressed as

$$P_{b,i}^{[k]} = \frac{2^{\gamma_{b,i}^{(\text{th})}} - 1}{\mathbf{d}_{C,b,i}^H \left(\mathcal{F}_C(\mathbf{p}_b^{[k-1]}, \mathbf{p}_E^{(t-1)}) - \mathbf{d}_{C,b,i}\mathbf{d}_{C,b,i}^H P_{b,i}^{[k-1]}\right)^{-1} \mathbf{d}_{C,b,i}}, \quad \text{for } 1 \leq i \leq K_{C,b}, \tag{43}$$

where k is the index of inner iteration for the power control algorithm. In the above iterative scheme, the initial transmit power is the optimized power of the $(t-1)$-th iteration, i.e., $\{P_{b,i}^{(t-1)}\}$. The optimized value of the t-th iteration, i.e., $\{P_{b,i}^{(t)}\}$, is set as the convergent value of (43).

Proposition 3. *If the proposed power control scheme has a stationary point, the stationary point is unique.*

Proof: Suppose $\mathbf{q}_1$ and $\mathbf{q}_2$ are two distinct stationary points of the proposed power control algorithm with $\mathbf{q}_i \succeq \mathbf{0}$ for $i = 1, 2$. Without loss of generality, we assume that there exist some elements in $\mathbf{q}_1$ larger than the corresponding elements in $\mathbf{q}_2$. We define

$$j^* = \arg\max_j \{[\mathbf{q}_1]_j / [\mathbf{q}_2]_j\}, \tag{44}$$

$$\xi = [\mathbf{q}_1]_{j^*} / [\mathbf{q}_2]_{j^*} > 1. \tag{45}$$

Then, we can get

$$\xi \mathbf{q}_2 \succeq \mathbf{q}_1. \tag{46}$$

Accordingly, we have

$$\begin{aligned}[\mathbf{q}_1]_{j^*} =& \frac{2^{\gamma_{b,i}^{(\mathrm{th})}}-1}{\mathbf{d}_{C,b,j^*}^H\left(\mathcal{F}_C(\mathbf{q}_1,\mathbf{p}_E^{(t-1)})-\mathbf{d}_{C,b,j^*}\mathbf{d}_{C,b,j^*}^H[\mathbf{q}_1]_{j^*}\right)^{-1}\mathbf{d}_{C,b,j^*}} \\ \leq& \frac{2^{\gamma_{b,i}^{(\mathrm{th})}}-1}{\mathbf{d}_{C,b,j^*}^H\left(\mathcal{F}_C(\xi\mathbf{p}_2,\mathbf{p}_E^{(t-1)})-\xi\mathbf{d}_{C,b,j^*}\mathbf{d}_{C,b,j^*}^H[\mathbf{q}_2]_{j^*}\right)^{-1}\mathbf{d}_{C,b,j^*}} \\ <& \frac{\xi(2^{\gamma_{b,i}^{(\mathrm{th})}}-1)}{\mathbf{d}_{C,b,j^*}^H\left(\mathcal{F}_C(\mathbf{p}_2,\mathbf{p}_E^{(t-1)})-\mathbf{d}_{C,b,j^*}\mathbf{d}_{C,b,j^*}^H[\mathbf{q}_2]_{j^*}\right)^{-1}\mathbf{d}_{C,b,j^*}} \\ =& \xi[\mathbf{q}_2]_{j^*}, \end{aligned} \tag{47}$$

which makes a contradiction with (45). Thus, Proposition 3 is proved. ∎

Proposition 4. *If the optimization problem is feasible, for any initial transmit power vector $\mathbf{q}_0$, the proposed power control algorithm converges to the unique stationary point.*

Proof: The proof is shown in Appendix A. ∎

The power control scheme in the cell-edge group is designed based on the optimized power coefficients of the cell-center groups. Following a similar way, it is not difficult to obtain the optimal power control scheme of the cell-edge group, which is given as

$$P_{E,i}^{[k]} = \frac{2^{\gamma_{E,i}^{(\mathrm{th})}}-1}{\mathbf{c}_{E,i}^H\left(\mathcal{F}_E(\mathbf{p}_1^{(t)},\mathbf{p}_2^{(t)},\mathbf{p}_E^{[k-1]})-\mathbf{c}_{E,i}\mathbf{c}_{E,i}^H P_{E,i}^{[k-1]}\right)^{-1}\mathbf{c}_{E,i}}, \quad \text{for } 1 \leq i \leq K_E. \tag{48}$$

Here, given the phase shift matrix $\mathbf{\Phi}^{(t-1)}$, the transmit powers of the cell-center groups and the cell-edge group have been devised by using (43) and (48), respectively. Since the both schemes in (43) and (48) are implemented iteratively, the convergent values of (43) and (48) are set as the optimized transmit powers of the t-th outer iteration. After the transmit powers are achieved, the optimal equalizers for the users in the cell-center and the cell-edge groups, i.e., $\mathbf{t}_{b,i}^{(t)}$ and $\mathbf{t}_{E,i}^{(t)}$, are computed by (20) and (23) accordingly.

C. Sequential Phase Shift Optimization

In this subsection, given the transmit powers $\{P_{b',i}^{(t)}\}$ and the equalizers $\{\mathbf{t}_{b',i}^{(t)}\}$, we aim at optimizing the phase shift matrix $\mathbf{\Phi}$. Accordingly, the optimization problem is converted into

$$\text{(P5) Find : } \mathbf{\Phi} \tag{49}$$

$$\text{s.t. } \log_2(1+\eta_{b',i}) \geq \gamma_{b',i}^{(\mathrm{th})},\ b' \in \{1,2,E\}, \forall i, \tag{50}$$

$$\omega_n \in \mathcal{X},\ 1 \leq n \leq N_S. \tag{51}$$

To facilitate the design of the phase shift matrix, we should reorganize the SINR expressions, which are reexpressed as

$$\eta_{b,i} = \frac{|\mathbf{t}_{b,i}^{(t)H}\mathbf{d}_{C,b,i}|^2 P_{b,i}^{(t)}}{\sum\limits_{i'=1}^{K_E}|Y_{b,i,i'}+\mathbf{g}_{b,i,i'}^H\mathbf{w}|^2 P_{E,i'}^{(t)} + I_{b,i}},\ b \in \{1,2\}, \tag{52}$$

$$\eta_{E,i} = \frac{|Y_{E,i,i}+\mathbf{g}_{E,i,i}^H\mathbf{w}|^2 P_{E,i}^{(t)}}{\sum\limits_{i'=1,i'\neq i}^{K_E}|Y_{E,i,i'}+\mathbf{g}_{E,i,i'}^H\mathbf{w}|^2 P_{E,i'}^{(t)} + I_{E,i}}, \tag{53}$$

where

$$I_{b,i} = \sum_{i'=1,i'\neq i}^{K_{C,b}}|\mathbf{t}_{b,i}^{(t)H}\mathbf{d}_{C,b,i'}|^2 P_{b,i'}^{(t)} + \sigma^2\mathbf{t}_{b,i}^{(t)H}\mathbf{t}_{b,i}^{(t)}, \tag{54}$$

$$I_{E,i} = \sum_{b=1}^{2}\sum_{i'=1}^{K_{C,b}}|\mathbf{t}_{E,i}^{(t)H}\tilde{\mathbf{d}}_{C,b,i'}|^2\varepsilon_{b,i'}P_{b,i'}^{(t)} + \sigma^2\mathbf{t}_{E,i}^{(t)H}\mathbf{t}_{E,i}^{(t)}, \tag{55}$$

$$Y_{b,i,i'} = \mathbf{t}_{b,i}^{(t)H}\mathbf{d}_{E,b,i'}, \tag{56}$$

$$Y_{E,i,i'} = \mathbf{t}_{E,i}^{(t)H}\left[\mathbf{d}_{E,1,i'}^T, \mathbf{d}_{E,2,i'}^T\right]^T, \tag{57}$$

$$\mathbf{g}_{b,i,i'} = \left(\mathbf{t}_{b,i}^{(t)H}\mathbf{G}_{b,S}\cdot\mathrm{diag}\{\mathbf{g}_{S,i'}\}\right)^H, \tag{58}$$

$$\mathbf{g}_{E,i,i'} = \left(\mathbf{t}_{E,i}^{(t)H}\begin{bmatrix}\mathbf{G}_{1,S}\\ \mathbf{G}_{2,S}\end{bmatrix}\cdot\mathrm{diag}\{\mathbf{g}_{S,i'}\}\right)^H, \tag{59}$$

$$\mathbf{w} = \mathrm{diag}\{\mathbf{\Phi}\} = [e^{\jmath\omega_1},\cdots,e^{\jmath\omega_{N_S}}]^T. \tag{60}$$

It is shown that there is no objective function in problem (P5). In order to reap an more efficient search path for the next power optimization subproblem, an objective function is inserted to improve the sum of all users' SINRs.[6] As a result, for satisfying the rate threshold of each user, the required transmit power is reduced in the sequential power optimization. It will be shown that the required transmit power has a rapid descent at the beginning of each outer iteration in the next section. Then, the optimization problem (P5) is reformulated as

$$\text{(P6) } \max_{\mathbf{\Phi}}: \sum_{b=1}^{2}\sum_{i=1}^{K_{C,b}}\eta_{b,i} + \sum\nolimits_{i=1}^{K_E}\eta_{E,i} \tag{61}$$

$$\text{s.t. } \eta_{b',i} \geq 2^{\gamma_{b',i}^{(\mathrm{th})}}-1,\ b' \in \{1,2,E\}, \forall i, \tag{62}$$

$$\omega_n \in \mathcal{X},\ 1 \leq n \leq N_S. \tag{63}$$

In this work, a sequential phase rotation approach is developed to solve this problem. To be specific, the new phase shift vector is obtained by rotating the phase shifts at the RIS sequentially. Following this concept, the new phase shift vector is achieved by an iteration scheme, i.e., the relationship between the k-th and the $(k-1)$-th phase shift vectors is given by

$$\mathbf{w}^{[k]} = \mathbf{\Gamma}_{N_S}\times\cdots\times\mathbf{\Gamma}_1\mathbf{w}^{[k-1]}, \tag{64}$$

where $\mathbf{\Gamma}_n$ is the rotation matrix for the n-th phase shift, which is expressed as $\mathbf{\Gamma}_n = \mathrm{diag}\{\mathbf{1}_{1\times(n-1)}, e^{\jmath\psi_n^{[k]}}, \mathbf{1}_{1\times(N_S-n)}\}$. It can be easily examined that the rotated phase shift $(\mathrm{mod}2\pi)$ still belongs to the sex $\mathcal{X}$ if $\psi_n^{[k]} \in \mathcal{X}$. Besides, defining $\mathbf{w}_n^{[k]} = \mathbf{\Gamma}_n\times\cdots\times\mathbf{\Gamma}_1\mathbf{w}^{[k-1]}$, we have

$$\mathbf{w}_n^{[k]} = \mathbf{\Gamma}_n\mathbf{w}_{n-1}^{[k]}. \tag{65}$$

It is clear that the design of $\mathbf{w}_n^{[k]}$ can be implemented based on $\mathbf{w}_{n-1}^{[k]}$. In the proposed scheme, the phase rotation matrices

[6]The problem (P5) can also be transformed to maximize the minimum SINR by using the max-min criterion, which can be addressed by invoking the combination of the bisection method and the proposed sequential phase shift optimization algorithm. Due to the limitation of space, the transformation by using the max-min criterion will be investigated in our future work.

$\{\mathbf{\Gamma}_n\}$ are designed in sequence. The optimization of the n-th matrix $\mathbf{\Gamma}_n$ is based on the last rotated value, i.e., $\mathbf{w}_{n-1}^{[k]}$. As a result, the SINRs for the users in the cell-center and the cell-edge groups are rewritten as

$$\eta_{b,i} = \frac{|\mathbf{t}_{b,i}^{(t)H}\mathbf{d}_{C,b,i}|^2 P_{b,i}^{(t)}}{\sum\limits_{i'=1}^{K_E} \mathcal{H}_{b,i,i',n}(\psi_n^{[k]}) + I_{b,i}},\ b \in \{1,2\}, \tag{66}$$

$$\eta_{E,i} = \frac{\mathcal{H}_{E,i,i,n}(\psi_n^{[k]})}{\sum\limits_{i'=1,i'\neq i}^{K_E} \mathcal{H}_{E,i,i',n}(\psi_n^{[k]}) + I_{E,i}}, \tag{67}$$

with

$$\begin{aligned}\mathcal{H}_{b',i,i',n}(\psi_n^{[k]}) =& \mu_{b',i,i',n}^{(0)} + \mu_{b',i,i',n}^{(1)}\cos\psi_n^{[k]} \\ &+ \mu_{b',i,i',n}^{(2)}\sin\psi_n^{[k]},\end{aligned} \tag{68}$$

$$\rho_{b',i,i',n} = Y_{b',i,i'} + \sum\nolimits_{n'=1,n'\neq n}^{N_S} [\mathbf{g}_{b',i,i'}]_{n'}^{*}[\mathbf{w}_{n-1}^{[k]}]_{n'}, \tag{69}$$

$$\xi_{b',i,i',n} = [\mathbf{g}_{b',i,i'}]_n^{*}[\mathbf{w}_{n-1}^{[k]}]_n, \tag{70}$$

$$\mu_{b',i,i',n}^{(0)} = (\rho_{b',i,i',n}^H \rho_{b',i,i',n} + \xi_{b',i,i',n}^H \xi_{b',i,i',n}) P_{E,i'}^{(t)}, \tag{71}$$

$$\mu_{b',i,i',n}^{(1)} = 2\Re\{\xi_{b',i,i',n}^H \rho_{b',i,i',n}\} P_{E,i'}^{(t)}, \tag{72}$$

$$\mu_{b',i,i',n}^{(2)} = 2\Im\{\xi_{b',i,i',n}^H \rho_{b',i,i',n}\} P_{E,i'}^{(t)}, \tag{73}$$

where $\Re\{\cdot\}$ and $\Im\{\cdot\}$ are used to extract the real and the imaginary parts of a complex number, respectively. In a further way, the problem (P6) is converted into a series of optimization subproblems of $\{\psi_n^{[k]}\}$, which is given by

$$\text{(P7)} \max_{\psi_n^{[k]}}: \ \mathcal{P}(\psi_n^{[k]}) \triangleq \sum_{b=1}^{2}\sum_{i=1}^{K_{C,b}} \eta_{b,i} + \sum_{i=1}^{K_E} \eta_{E,i} \tag{74}$$

$$\text{s.t. } \mathcal{Y}_{b,i,n}(\psi_n^{[k]}) \geq (2^{\gamma_{b,i}^{(\text{th})}} - 1) I_{b,i}, \forall b, i, \tag{75}$$

$$\mathcal{Y}_{E,i,n}(\psi_n^{[k]}) \geq (2^{\gamma_{E,i}^{(\text{th})}} - 1) I_{E,i}, \forall i, \tag{76}$$

$$\psi_n^{[k]} \in \mathcal{X}, \tag{77}$$

where

$$\begin{aligned}\mathcal{Y}_{b,i,n}(\psi_n^{[k]}) =& |\mathbf{t}_{b,i}^{(t)H}\mathbf{d}_{C,b,i}|^2 P_{b,i}^{(t)} \\ &- (2^{\gamma_{b,i}^{(\text{th})}} - 1)\sum_{i'=1}^{K_E} \mathcal{H}_{b,i,i',n}(\psi_n^{[k]}),\end{aligned} \tag{78}$$

$$\begin{aligned}\mathcal{Y}_{E,i,n}(\psi_n^{[k]}) =& \mathcal{H}_{E,i,i,n}(\psi_n^{[k]}) \\ &- (2^{\gamma_{E,i}^{(\text{th})}} - 1) \sum_{i'=1,i'\neq i}^{K_E} \mathcal{H}_{E,i,i',n}(\psi_n^{[k]}).\end{aligned} \tag{79}$$

It is shown that in problem (P7), there is only one optimization variable. By employing the penalty method, the n-th optimal rotation value in the k-th inner iteration is achieved by

$$\begin{aligned}\psi_n^{[k]\star} = \arg \max_{\psi_n^{[k]}\in\mathcal{X}} \Big\{ &\mathcal{P}(\psi_n^{[k]}) \\ &+ \sum_{b'=\{1,2,E\}} \sum_{\forall i} \Delta\big\{\mathcal{Y}_{b',i,n}(\psi_n^{[k]}) - (2^{\gamma_{b',i}^{(\text{th})}} - 1) I_{b',i}\big\}\Big\},\end{aligned} \tag{80}$$

where $\Delta\{x\}$ is an index function given by $\Delta\{x\} = \begin{cases} 0, & x \geq 0 \\ -M, & \text{otherwise} \end{cases}$, with M denoting the penalty parameter. Due to the discrete features of phase shifts, the optimal solution of phase shift is obtained by the one-dimension search.

Because the objective value of problem (P7) is non-decreased with the rotation indices and has an upper bound, the proposed phase rotation algorithm is convergent accordingly. The optimal rotation value of the n-th phase shift is: $\psi_n^{\star} = \sum_{k=1}^{K} \psi_n^{[k]\star}$, where K denotes the number of inner iterations when the objective value of problem (P7) converges. When the optimal rotation value is achieved, the optimal phase shift can be computed accordingly, which is given by

$$\omega_n^{(t)} = \left(\omega_n^{(t-1)} + \psi_n^{\star}\right) \mod 2\pi,\ \forall n. \tag{81}$$

As a result, in the t-th iteration, the optimal phase shift matrix is constructed by

$$\mathbf{\Phi}^{(t)} = \text{diag}\left\{\omega_1^{(t)}, \omega_2^{(t)}, \cdots, \omega_{N_S}^{(t)}\right\}. \tag{82}$$

D. Design Summary and Several Discussions

In the above two subsections, the transmit power and the phase shifts in the t-th outer iteration are optimized by two iterative methods, respectively. The implementation procedures of the overall scheme are presented in Algorithm 1.

Algorithm 1 The proposed power control and phase shift scheme

Set the initial transmit power $\{P_{b',i}^{(0)}\}$, the initial phase shifts $\{\omega_n^{(0)}\}$, the termination thresholds of the outer and inner iterations ζ_{out} and ζ_{in}, and $t = 0$;
repeat
  $t = t + 1$;
  repeat
    Update the transmit power of the cell-center groups in parallel by using (43);
    Update the transmit power of the cell-edge group in parallel by using (48);
  until $|\sum_{i=1}^{K_{C,b}(K_E)} P_{b',i}^{[k]} - \sum_{i=1}^{K_{C,b}(K_E)} P_{b',i}^{[k-1]}| < \zeta_{\text{in}}$
  Set the optimal transmit power in the t-th outer iteration, i.e, $P_{b,i}^{(t)}$ and $P_{E,i}^{(t)}$, as the convergent values of (43) and (48);
  Compute the optimal equalizer, i.e., $\mathbf{t}_{b,i}^{(t)}$ and $\mathbf{t}_{E,i}^{(t)}$, by using (20) and (23);
  repeat
    Update each phase shifter in sequence by using (80);
  until $\sum_{n=1}^{N_S} \psi_n^{[k]} = 0$
  Compute the optimal phase shift matrix in the t-th outer iteration, i.e., $\mathbf{\Phi}^{(t)}$, by using (81) and (82);
until $|\sum_{b=1}^{2}\sum_{i=1}^{K_{C,b}}(P_{b,i}^{(t)} - P_{b,i}^{(t-1)}) + \sum_{i=1}^{K_E}(P_{E,i}^{(t)} - P_{E,i}^{(t-1)})| < \zeta_{\text{out}}$.

Remark 1 [The convergence of the overall algorithm]: It has been demonstrated that each user's transmit power is

This article has been accepted for publication in IEEE Transactions on Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2022.3177773

convergent in the phase of power coefficient updating. After the convergence of power control schemes, we have

$$\eta_{b',i} = 2^{\gamma^{(\text{th})}_{b',i}} - 1,\ b' \in \{1, 2, E\}. \tag{83}$$

According to problem (P6), the sum of SINRs is increased in the optimization of phase shifts, while each user's SINR is not decreased due to the constraint (62). Then, after the implementation of phase shift optimization, we have

$$\eta_{b',i} \geq 2^{\gamma^{(\text{th})}_{b',i}} - 1,\ b' \in \{1, 2, E\}. \tag{84}$$

In the t-th outer iteration, the initial transmit power is set as the convergence value of the $(t-1)$-th iteration, i.e., $P^{[0]}_{b',i} = P^{(t-1)}_{b',i}$. Thus, in the next stage of power control, the SINR of the cell-center users is given by

$$\begin{aligned}\eta_{b,i} =& \mathbf{d}^H_{C,b,i}\left(\mathcal{F}_{C,b}(\mathbf{p}^{[0]}_b, \mathbf{p}^{(t-1)}_E) - \mathbf{d}_{C,b,i}\mathbf{d}^H_{C,b,i}P^{[0]}_{b,i}\right)^{-1} \\ &\times \mathbf{d}_{C,b,i}P^{[0]}_{b,i} \\ \geq& 2^{\gamma^{(\text{th})}_{b,i}} - 1,\ b \in \{1, 2\}.\end{aligned} \tag{85}$$

According to the power control scheme in (43), we further have

$$\begin{aligned}P^{[1]}_{b,i} =& \frac{2^{\gamma^{(\text{th})}_{b,i}} - 1}{\mathbf{d}^H_{C,b,i}\left(\mathcal{F}_{C,b}(\mathbf{p}^{[0]}_b, \mathbf{p}^{(t-1)}_E) - \mathbf{d}_{C,b,i}\mathbf{d}^H_{C,b,i}P^{[0]}_{b,i}\right)^{-1}\mathbf{d}_{C,b,i}} \\ \leq& P^{[0]}_{b,i},\ b \in \{1, 2\}.\end{aligned} \tag{86}$$

Here, it is proved that after the design of phase shifts, the transmit power at the beginning of the t-th outer iteration is no more than the convergence value of the $(t-1)$-th iteration, which will also be demonstrated by Fig. 2 in Section IV. Based on the derivations in Appendix A, we can further get that the transmit power $\{P^{[k]}_{b,i}\}$ converges to the stationary point. Accordingly, the convergence of the cell-edge group's transmit power can be derived in the same way. Thus, the convergence of the objective function value (i.e., the sum of all users' transmit power) of the proposed algorithm is guaranteed.

Remark 2 [Computation complexity]: In the phase of transmit power and equalizer designs, the computational complexity is dominant by the matrix multiplications and the matrix inversions in (43), (48), (20), and (23). Thus, the computation complexity is about $\mathcal{O}(\max\{N_S N_A \cdot K_E, N_A^3 \cdot (\sum_{b=1}^{2} K_{C,b} + K_E)\})$. In the phase of phase shifts, the computation complexity is about $\mathcal{O}(2^X N_S^2 \cdot (\sum_{b=1}^{2} K_{C,b} + K_E))$. Due to the low expense of RIS, the number of reflecting elements is larger than that of receiving antennas at the BS generally. In such case, the computation complexity of the overall algorithm in each outer iteration is $\mathcal{O}(\max\{2^X N_S^2, N_A^3\} \cdot (\sum_{b=1}^{2} K_{C,b} + K_E))$. In practice, the number of resolution bits, X, is equal to that of positive-intrinsic negative (PIN) diodes for each reflection unit. Due to the limitation of volume and hardware complexity, the number of resolution bits is not large in the practical systems. Thus, it will not entail a high computational complexity for the proposed phase shift scheme. In current prototype systems, the number of resolution bits for each reflecting element is 1 or 2 generally.

Remark 3 [System overhead]: In the proposed scheme, the channel state information (CSI) between the users and the BSs should be available at the BSs or the CU, which can be obtained by channel estimation methods for RIS-aided communication systems [40]. After the design procedures are completed, each BS broadcasts the optimal transmit power coefficients to its associated users, whilst the optimized phase shifts are delivered to the RIS via the RIS controller.

TABLE I
SIMULATION PARAMETERS

Parameters	Values
The path loss exponents $[\beta_{S,U}, \beta_{B,S}, \beta_{B,U}]$	[2.5, 3, 3.5]
The Rician-K factors $[\kappa_{S,U}, \kappa_{B,S}]$	[2, 2]
The coverage radius of the BS R_B	500 m
The number of antennas at the BS N_A	4
The radius of the cell-center region R_C	200 m
The coverage radius of the RIS R_S	50 m
The number of resolution bits X	4
The number of users per group $[K_{C,b}, K_E]$	[4, 4]
The number of reflecting elements N_S	$20 \sim 80$
The termination thresholds $[\zeta_{\text{in}}, \zeta_{\text{out}}]$	$[10^{-5}, 10^{-5}]$
The minimum rate requirement per user $\gamma^{(\text{th})}_{b',i}$	$1 \sim 3$ b/s/Hz
The error propagation factor $\varepsilon_{b,i}$	$0.001 \sim 0.1$
The maximum power budget $P_{\max}$	30 dBm
The noise power σ^2	-99 dBm

IV. SIMULATION RESULTS

In this section, the total power consumption of the proposed scheme is evaluated. The simulation setups are given as follows. Each element in spatial correlation matrices is given by $[\mathbf{\Upsilon}_{B,R}]_{x_1,x_2} = 0.5^{|x_2 - x_1|}$, $[\mathbf{\Upsilon}_{S,R}]_{x_1,x_2} = 0.5^{|x_2 - x_1|}$, and $[\mathbf{\Upsilon}_{S,T}]_{x_1,x_2} = 0.5^{|x_2 - x_1|}$. The LoS component for the channel is characterized by [38]

$$\bar{\mathbf{F}} = \mathbf{h}_{N_r}(\theta^{\text{AoA}}) \times \mathbf{h}^H_{N_t}(\theta^{\text{AoD}}), \tag{87}$$

with

$$\mathbf{h}_{N_r}(\theta^{\text{AoA}}) = [1, e^{\jmath 2\pi \frac{d}{\lambda}\sin(\theta^{\text{AoA}})}, \cdots, e^{\jmath 2\pi \frac{d(N_r-1)}{\lambda}\sin(\theta^{\text{AoA}})}]^T, \tag{88}$$

$$\mathbf{h}_{N_t}(\theta^{\text{AoD}}) = [1, e^{\jmath 2\pi \frac{d}{\lambda}\sin(\theta^{\text{AoD}})}, \cdots, e^{\jmath 2\pi \frac{d(N_t-1)}{\lambda}\sin(\theta^{\text{AoD}})}]^T, \tag{89}$$

where θ^{AoA} and θ^{AoD} are the angles of arrival and departure, respectively, which are randomly selected within the region $[0, 2\pi)$, N_r and N_t are the number of antennas/elements at the receiver and the transmitter, respectively, λ and d are respectively the wavelength and the distances of the antenna/element separation, with $d/\lambda = 1/2$. Unless otherwise specified, the other simulation parameters are tabulated in Table I. The initial phase shifts are generated randomly within the set $\mathcal{X}$. The initial transmit power of each user is set as $P^{(0)}_{b',i} = 20$ dBm for $b' \in \{1, 2, E\}$.

For comparison, several benchmark schemes are introduced. *CoMP-NOMA w/ RPS:* The optimal power control is adopted by both the cell-center and the CoMP groups, whilst each phase shifter at the RIS is chosen randomly within the set $\mathcal{X}$. *RIS-OMA w/ PC:* The cell-center and the CoMP groups are organized on the orthogonal and equal time slots. Besides, the optimal power control in [39] is adopted by the users in each group. *CoMP-NOMA w/o RIS:* The cell-center and the CoMP groups are scheduled by the BSs under the NOMA mode.

This article has been accepted for publication in IEEE Transactions on Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2022.3177773

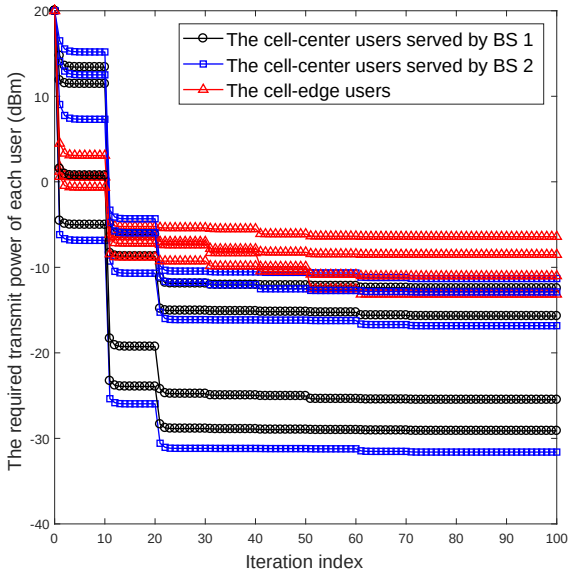


Fig. 2. Each user's transmit power requirement versus the iteration index with $\gamma_{b',i}^{(\mathrm{th})} = 2$ b/s/Hz and $\varepsilon_{b,i} = 0.001$.

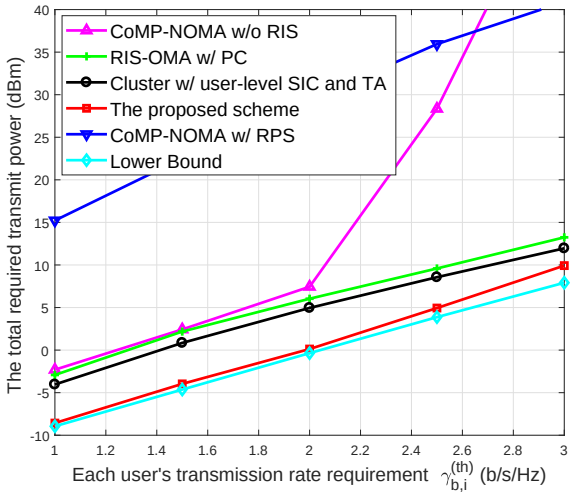


Fig. 3. The total required transmit power versus the transmission rate requirement with $N_S = 50$ and $\varepsilon_{b,i} = 0.001$.

However, the transmission of the CoMP group is not assisted by the RIS. *Cluster w/ user-level SIC and TA* [22]: The active users are divided into multiple clusters. Different clusters are scheduled on the orthogonal time slots by following OMA principle, while the users in the same cluster are decoded sequentially with the aid of user-level SIC by using NOMA principle. Besides, the time allocation (TA) for each cluster is optimized to minimize the total power consumption. In addition, we provide the scheme that the RIS can adjust phase shifts continuously as the *lower bound* of the performance.

A. The Convergence Performance

Fig. 2 plots the transmit power of each user versus the number of iterations for one simulation realization. For convenience, the number of the inner and the outer iterations are set as 10 and 10, respectively. It is observed that the required transmit power of each user in the cell-center or the cell-edge groups is convergent. On one hand, the transmit power in the inner iteration decreases with the iteration indices. On the other hand, there is a rapid decline at the begin of each outer iteration. The reason is because the phase shifts are devised for maximizing the sum of all users' SINRs, which provides a more efficient search path for optimizing the transmit power in the following iterations. Since the required transmit power per user is convergent, it is easy to obtain that the objective function value (i.e., the total transmit power) of the proposed algorithm is convergent accordingly, which agrees with the discussion in Remark 1.

B. The Required Transmit Power Consumption

Fig. 3 plots the required transmit power for different transmission rate requirements. It is noted that as the transmission rate increases, the total transmit power is increased for all schemes. It is observed that the total power consumption of the proposed scheme is quite close to the lower bound, especially for low transmission rate requirements. Besides, the required power consumption of the proposed scheme is less than various benchmark schemes. The reasons are given as follows. Compared with the CoMP-NOMA w/ RPS, the benefit of our proposed scheme originates from the fact that the transmit power, the phase shifts, and the equalizers are collaboratively designed to cut down the total power consumption. Compared with the RIS-OMA w/ PC, the transmit power is decreased because of the advantage of NOMA over OMA. The reason for the phenomenon that the proposed scheme is superior to the CoMP-NOMA w/o RIS is given as follows. The deployment of RIS has the capability of enhancing the received power levels of the cell-edge users. As a result, the required transmit power is reduced for satisfying the predefined transmission rate per user. What's more, our proposed scheme outperforms the Cluster w/ user-level SIC and TA because the new concept of GSIC is leveraged to cope with the inter-group interference. In the conventional user-level SIC framework, the different clusters are served by following the OMA mode, while the users in the same cluster are decoded by using the user-level SIC. This approach has shortcomings of degraded power efficiency because the transceivers are not jointly designed and the inter-group coordination is not considered. However, in the proposed GSIC framework, the users are divided into multiple groups according to the path losses. The users' signals in each group are decoded in parallel and subtracted simultaneously by invoking GSIC. The benefit of the proposed scheme is that the transceivers in each group are designed jointly by considering intra-group and inter-group interference, and thus the power efficiency is improved.

C. The Effect of the Number of Reflecting Elements

Fig. 4 plots the total transmit power for different number of reflecting elements. As expected, the power consumption is decreased with the increasing number of reflecting units for the proposed scheme. This is because the cascaded channel is strengthened as the number of reflecting elements

This article has been accepted for publication in IEEE Transactions on Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2022.3177773

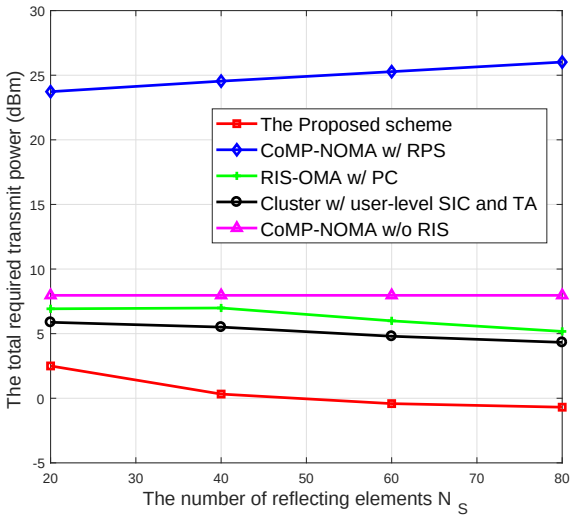


Fig. 4. The total required transmit power versus the number of reflecting elements with $\gamma_{b',i}^{(\mathrm{th})} = 2$ b/s/Hz and $\varepsilon_{b,i} = 0.001$.

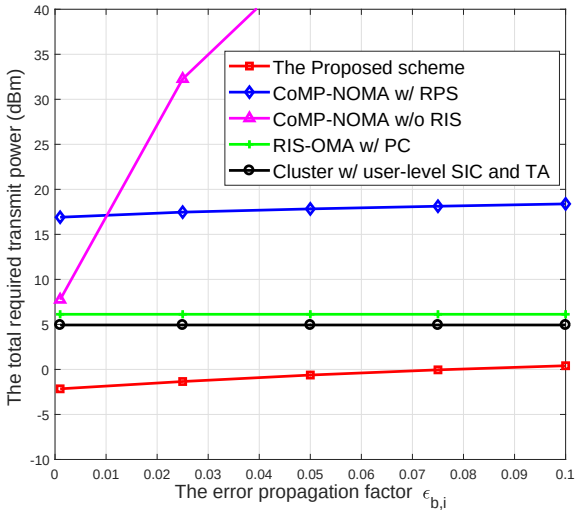


Fig. 5. The total required transmit power versus the error propagation factor with $N_S = 50$, $K_{C,b} = K_E = 3$, and $\gamma_{b',i}^{(\mathrm{th})} = 2$ b/s/Hz.

increases. However, for the CoMP-NOMA w/ RPS scheme, the required transmit power is increased slightly with the number of reflecting elements, which is different from the three schemes with optimized phase shifts. It is because the cascaded channel may neutralize the equivalent channel gain by using the random phase shifts. It reflects the importance of optimizing the phase shifts for multi-user multi-antenna RIS-NOMA networks. Besides, it is seen that the proposed scheme is superior to the four benchmarking algorithms for different RIS configurations. The reason has been presented in our forgoing analyses.

D. The Impact of Error Propagation Factor

Fig. 5 plots the total transmit power for different error propagation factors. As the error propagation factor increases, the performance of total power consumption is degraded for the three schemes, i.e., the proposed scheme, the CoMP-NOMA w/ RPS, and the CoMP-NOMA w/o RIS. The reason is that, with the increase of the error propagation factor, the residual inter-group interference caused the cell-center groups is increased. Besides, compared with the proposed RIS-aided NOMA scheme, it is noted that the required transmit power of the CoMP-NOMA w/o RIS grows rapidly with the error propagation factor. It indicates the effectiveness of RIS in dealing with the residual interference. However, for the RIS-OMA w/ PC and the Cluster w/ user-level SIC and TA, the transmit power keeps unchanged almost with the growth of error propagation factor because there doesn't exist the residual inter-group interference. It implies, compared with the RIS-OMA w/ PC and the Cluster w/ user-level SIC and TA, the performance gain of the proposed method fades gradually as the error propagation factor increases. It is as expected that when the error propagation factor is greater than a threshold, the Cluster w/ user-level SIC and TA scheme will outperform the proposed method in terms of transmit power consumption. Fortunately, the error propagation factor is not large in advanced communication systems.

E. The Influence of the Number of Resolution Bits

Fig. 6 plots the total transmit power consumption versus the number of resolution bits. It is observed that the proposed algorithm is superior to the four benchmark schemes for the different number of resolution bits. Besides, it is noted that, with the increasing number of resolution bits, the total power consumption of the proposed scheme is reduced gradually. The reason is given as follows. As the number of resolution bits increases, the feasible set for the considered optimization problem is expanded. Thus, more candidate phase shifters can be picked to manipulate the cell-edge user signals' phases with the dual goals, i.e., useful signal enhancement and interference suppression. However, these benefits are achieved at the cost of improving circuit complexity of the RIS. Besides, for the CoMP-NOMA w/ RPS scheme, the required transmit power keeps unchanged almost as the number of resolution bits becomes large. It reveals the significance of optimizing the phase shifts in the multi-cell RIS-NOMA systems.

V. Conclusion

In this work, both RIS and CoMP were leveraged to boost the cell-edge users' performance for multi-cell NOMA networks. We investigated the power minimization problem for RIS-empowered multi-group CoMP-NOMA networks with error propagation. Particularly, the GSIC was conceived for removing inter-group interference, whilst the power control and the equalizers were jointly designed to cope with the remaining interference. By using the relationship between the optimal equalizer and the transmit power, the original problem was converted into a joint optimization problem about power control and phase shifts. To further render the transformed problem tractable, an alternating method was proposed to devise the transmit power and the phase shifts. To proceed, a parallel iteration algorithm was proposed to update the users' transmit power for each group, and a sequential rotation algorithm was developed to optimize the phase shifts in sequence. Extensive simulation results demonstrated that the

This article has been accepted for publication in IEEE Transactions on Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2022.3177773

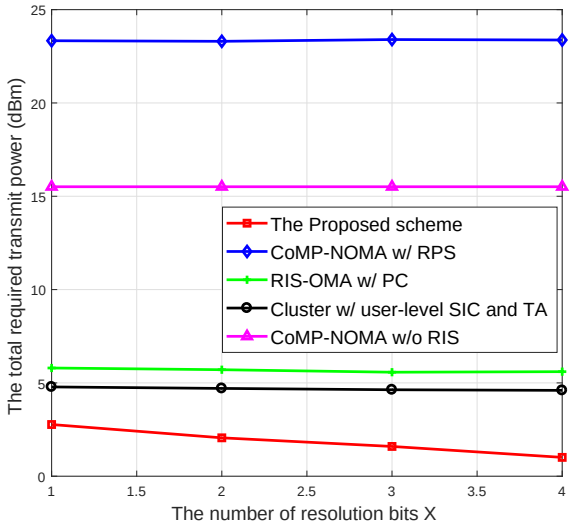


Fig. 6. The total required transmit power versus the number of resolution bits with $N_S = 50$, $\gamma_{b',i}^{(\text{th})} = 2$ b/s/Hz, and $\varepsilon_{b,i} = 0.001$.

total transmit power consumption of the proposed algorithm is less than those of various benchmarks.

Appendix A
The Proof of Proposition 4

Before proceeding to the proof of Proposition 4, we present a useful property at first. For two initialization values $\mathbf{q}_1$ and $\mathbf{q}_2$ with $\mathbf{q}_2 \succeq \mathbf{q}_1$, we can get that, in the iteration of the proposed power control algorithm, the entries in the vectors satisfy the following relationship:

$$[\mathbf{q}_2^{[k]}]_i \geq [\mathbf{q}_1^{[k]}]_i, \ \forall i, \ k = 1, 2, \cdots, \tag{90}$$

where k is the index of the inner iteration. Therefore, we have

$$\mathbf{q}_2^{[k]} \succeq \mathbf{q}_1^{[k]}, \ k = 1, 2, \cdots. \tag{91}$$

Suppose $\mathbf{q}^*$ is the stationary point of the proposed power control algorithm. In the following, we will prove Proposition 4 by considering the following two cases.

Case 1: $\mathbf{q}_0 \prec \mathbf{q}^*$. It can be obtained that $\mathbf{q}_0$ is not a feasible point for the optimization problem; otherwise, it makes a contradiction with the assumption that $\mathbf{q}^*$ is the unique stationary point. Then, we have

$$\begin{aligned}&\mathbf{d}_{C,b,i}^H(\mathcal{F}_C(\mathbf{q}_0, \mathbf{p}_E^{(t-1)}) - [\mathbf{q}_0]_i \mathbf{d}_{C,b,i}\mathbf{d}_{C,b,i}^H)^{-1}\mathbf{d}_{C,b,i}[\mathbf{q}_0]_i \\ <&2^{\gamma_{b,i}^{(\text{th})}} - 1, \ \forall i.\end{aligned} \tag{92}$$

By employing the proposed power updating algorithm, we can get

$$\begin{aligned}[\mathbf{q}_0^{[1]}]_i =& \frac{2^{\gamma_{b,i}^{(\text{th})}} - 1}{\mathbf{d}_{C,b,i}^H(\mathcal{F}_C(\mathbf{q}_0, \mathbf{p}_E^{(t-1)}) - [\mathbf{q}_0]_i \mathbf{d}_{C,b,i}\mathbf{d}_{C,b,i}^H)^{-1}\mathbf{d}_{C,b,i}} \\ >& [\mathbf{q}_0]_i, \ \forall i.\end{aligned} \tag{93}$$

In a similar way, we further have

$$\begin{aligned}[\mathbf{q}_0^{[2]}]_i =& \frac{2^{\gamma_{b,i}^{(\text{th})}} - 1}{\mathbf{d}_{C,b,i}^H(\mathcal{F}_C(\mathbf{q}_0^{[1]}, \mathbf{p}_E^{(t-1)}) - [\mathbf{q}_0^{[1]}]_i \mathbf{d}_{C,b,i}\mathbf{d}_{C,b,i}^H)^{-1}\mathbf{d}_{C,b,i}} \\ >& \frac{2^{\gamma_{b,i}^{(\text{th})}} - 1}{\mathbf{d}_{C,b,i}^H(\mathcal{F}_C(\mathbf{q}_0, \mathbf{p}_E^{(t-1)}) - [\mathbf{q}_0]_i \mathbf{d}_{C,b,i}\mathbf{d}_{C,b,i}^H)^{-1}\mathbf{d}_{C,b,i}} \\ =& [\mathbf{q}_0^{[1]}]_i.\end{aligned} \tag{94}$$

According to (93) and (94), by using the mathematical induction, we have

$$\mathbf{q}_0 \preceq \mathbf{q}_0^{[1]} \preceq \mathbf{q}_0^{[2]} \preceq \cdots \preceq \mathbf{q}_0^{[k]} \preceq \cdots. \tag{95}$$

Besides, based on the property in (91), we can get that

$$\mathbf{q}_0^{[k]} \preceq \mathbf{q}^*, k = 1, 2, \cdots. \tag{96}$$

Thus, it can be obtained that as k increases, $\mathbf{q}_0^{[k]}$ is non-decreased with the upper bound $\mathbf{q}^*$. Based on the monotone bounded theorem, we have that the sequences $\{\mathbf{q}_0^{[k]}\}$ are convergent. Recalling that the stationary point is unique in Proposition 3, we can get that the transmit power will converge to the unique stationary point $\mathbf{q}^*$.

Case 2: $\mathbf{q}_0 \not\prec \mathbf{q}^*$. In such case, we can find two constants $\delta_1 < 1$ and $\delta_2 > 1$ such that

$$\mathbf{q}_1 = \delta_1 \mathbf{q}^* \preceq \mathbf{q}_0 \preceq \mathbf{q}_2 = \delta_2 \mathbf{q}^*. \tag{97}$$

According to the property in (91), we have

$$\mathbf{q}_1^{[k]} \preceq \mathbf{q}_0^{[k]} \preceq \mathbf{q}_2^{[k]}, \ \forall k. \tag{98}$$

Based on the derivations in Case 1, we can get that

$$\lim_{k \to \infty} \mathbf{q}_1^{[k]} = \mathbf{q}^*. \tag{99}$$

Besides, we can obtain

$$\begin{aligned}[\mathbf{q}_2^{[1]}]_i =& \frac{2^{\gamma_{b,i}^{(\text{th})}} - 1}{\mathbf{d}_{C,b,i}^H(\mathcal{F}_C(\mathbf{q}_2, \mathbf{p}_E^{(t-1)}) - [\mathbf{q}_2]_i \mathbf{d}_{C,b,i}\mathbf{d}_{C,b,i}^H)^{-1}\mathbf{d}_{C,b,i}} \\ \leq& \delta_2 \frac{2^{\gamma_{b,i}^{(\text{th})}} - 1}{\mathbf{d}_{C,b,i}^H(\mathcal{F}_C(\mathbf{q}^*, \mathbf{p}_E^{(t-1)}) - [\mathbf{q}^*]_i \mathbf{d}_{C,b,i}\mathbf{d}_{C,b,i}^H)^{-1}\mathbf{d}_{C,b,i}} \\ =& \delta_2 [\mathbf{q}^*]_i \\ =& [\mathbf{q}_2]_i, \ \forall i.\end{aligned} \tag{100}$$

Accordingly, we further have

$$\begin{aligned}[\mathbf{q}_2^{[2]}]_i =& \frac{2^{\gamma_{b,i}^{(\text{th})}} - 1}{\mathbf{d}_{C,b,i}^H(\mathcal{F}_C(\mathbf{q}_2^{[1]}, \mathbf{p}_E^{(t-1)}) - [\mathbf{q}_2^{[1]}]_i \mathbf{d}_{C,b,i}\mathbf{d}_{C,b,i}^H)^{-1}\mathbf{d}_{C,b,i}} \\ \leq& \frac{2^{\gamma_{b,i}^{(\text{th})}} - 1}{\mathbf{d}_{C,b,i}^H(\mathcal{F}_C(\mathbf{q}_2, \mathbf{p}_E^{(t-1)}) - [\mathbf{q}_2]_i \mathbf{d}_{C,b,i}\mathbf{d}_{C,b,i}^H)^{-1}\mathbf{d}_{C,b,i}} \\ =& [\mathbf{q}_2^{[1]}]_i, \ \forall i.\end{aligned} \tag{101}$$

By invoking the mathematical induction, it can be proved that $\{[\mathbf{q}_2^{[k]}]_i\}$ are non-increased with the iteration index k, i.e.,

$$[\mathbf{q}_2]_i \geq [\mathbf{q}_2^{[1]}]_i \geq [\mathbf{q}_2^{[2]}]_i \geq \cdots \geq [\mathbf{q}_2^{[k]}]_i \geq \cdots. \tag{102}$$

Besides, it is clear that the transmit powers $\{[\mathbf{q}_2^{[k]}]_i\}$ are always larger than 0. Based on the monotone bounded theorem, it can be obtained that the sequences $\{\mathbf{q}_2^{[k]}\}$ are convergent.

Since $\mathbf{q}^*$ is the unique stationary point of the proposed iteration algorithm, it can be derived that

$$\lim_{k\to\infty} \mathbf{q}_2^{[k]} = \mathbf{q}^*. \tag{103}$$

According to (98), (99) and (103), By invoking the Squeeze theorem, we have

$$\lim_{k\to\infty} \mathbf{q}_0^{[k]} = \mathbf{q}^*. \tag{104}$$

Here, the proof of Proposition 4 is completed.

REFERENCES

[1] Y. Zhao, W. Zhai, J. Zhao, T. Zhang, S. Sun, D. Niyato, and K.-Y. Lam, "A comprehensive survey of 6G wireless communications," *arXiv:2101.03889v1*, Dec. 2020. [Online] Avaiable: https://arxiv.org/abs/2101.03889v1.

[2] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface aided wireless communications: A tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313-3351, May 2021.

[3] H. Wang, C. Liu, Z. Shi, Y. Fu, and R. Song, "On power minimization for IRS-aided downlink NOMA systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 11, pp. 1808-1811, Nov. 2020.

[4] 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, Nov. 2020.

[5] X. You *et al.*, "Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts," *Sci. China Inf. Sci.*, vol. 64, no. 1, p. 110301, Jan. 2021.

[6] Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Nonorthogonal multiple access for 5G and beyond," *Proc. IEEE*, vol. 105, no. 12, pp. 2347-2381, Dec. 2017.

[7] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. Bhargava, "A survey on non-orthogonal multiple access for 5G networks: Research challenges and future trends," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2181-2195, Oct. 2017.

[8] Y. Fu, Y. Chen, and C. W. Sung, "Distributed power control for the downlink of multi-cell NOMA systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6207-6220, Sept. 2017.

[9] A. S. D. Sena *et al.*, "What role do intelligent reflecting surfaces play in multi-antenna non-orthogonal multiple access?," *IEEE Wireless Commun.*, vol. 27, no. 5, pp. 23-31, Oct. 2020.

[10] M. Fu, Y. Zhou, and Y. Shi, "Intelligent reflecting surface for downlink non-orthogonal multiple access networks," *IEEE Globecom Workshops (GC Wkshps)*, Dec. 2019, pp. 1-6.

[11] X. Mu, Y. Liu, L. Guo, J. Lin, and N. Al-Dhahir, "Capacity and optimal resource allocation for IRS-assisted multi-user communication systems," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 3771-3786, Jun. 2021.

[12] Z. Ding, R. Schober, and H. V. Poor, "On the impact of phase shifting designs on IRS-NOMA," *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1596-1600, Oct. 2020.

[13] 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.

[14] J. Zuo, Y. Liu, Z. Qin, and N. Al-Dhahir, "Resource allocation in intelligent reflecting surface assisted NOMA systems," *IEEE Trans. Commun.*, vol. 68, no. 11, pp. 7170-7183, Nov. 2020.

[15] M. Elhattab, M. A. Arfaoui, C. Assi, and A. Ghrayeb, "RIS-assisted joint transmission in a two-cell downlink NOMA cellular system," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 4, pp. 1270-1286, Apr. 2022.

[16] M. Zeng, X. Li, G. Li, W. Hao, and O. A. Dobre, "Sum rate maximization for IRS-assisted uplink NOMA," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 234-238, Jan. 2021.

[17] Y. Liu, X. Mu, X. Liu, M. D. Renzo, Z. Ding, and R. Schober, "Reconfigurable intelligent surface (RIS) aided multi-user networks: Interplay between NOMA and RIS," *ArXiv:2011.13336v1*, 2020. [Online] Available: https://arxiv.org/abs/2011.13336v1.

[18] F. Fang, Y. Xu, Q.-V. Pham, and Z. Ding, "Energy-efficient design of IRS-NOMA networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 14088-14092, Nov. 2020.

[19] G. Yang, X. Xu, Y.-C. Liang, and M. D. Renzo, "Reconfigurable intelligent surface assisted non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 3137-3151, May 2021.

[20] Y. Li, M. Jiang, Q. Zhang, and J. Qin, "Joint beamforming design in multi-cluster MISO NOMA reconfigurable intelligent surface-aided downlink communication networks," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 664-674, Jan. 2021.

[21] X. Xie, F. Fang, and Z. Ding, "Joint optimization of beamforming, phase-shifting and power allocation in a multi-cluster IRS-NOMA network," *IEEE Trans. Veh. Technol.*, vol. 70, no. 8, pp. 7705-7717, Jun. 2020.

[22] D. Zhang, Q. Wu, M. Cui, G. Zhang, and D. Niyato, "Throughput maximization for IRS-assisted wireless powered hybrid NOMA and TDMA," *IEEE Wireless Commun. Lett.*, vol. 10, no. 9, pp. 1944-1948, Sept. 2021.

[23] M. Elhattab, M.-A. Arfaoui, and C. Assi, "CoMP transmission in downlink NOMA-based heterogeneous cloud radio access networks," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7779-7794, Dec. 2020.

[24] Y. Sun, Z. Ding, X. Dai, and O. A. Dobre, "On the performance of network NOMA in uplink CoMP systems: A stochastic geometry approach," *IEEE Trans. Commun.*, vol. 67, no. 7, pp. 5084-5098, Jul. 2019.

[25] W. Wang, L. Yang, A. Meng, Y. Zhan, D. W. K. Ng, "Resource allocation for IRS-aided JP-CoMP downlink cellular networks with underlaying D2D communications," *IEEE Trans. Wireless Commun.*, early access, pp. 1-15, 2021.

[26] A. Rezaei, P. Azmi, N. M. Yamchi, M. R. Javan, and H. Yanikomeroglu, "Robust resource allocation for cooperative MISO-NOMA-based heterogeneous networks," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 3864-3878, Jun. 2021.

[27] C. Pan, H. Ren, K. Wang, W. Xu, M. Elkashlan, A. Nallanathan, and L. Hanzo, "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5218-5233, Aug. 2020.

[28] H. Wang, C. Liu, Z. Shi, Y. Fu, and R. Song, "Power minimization for two-cell IRS-aided NOMA systems with joint detection," *IEEE Commun. Lett.*, vol. 25, no. 5, pp. 1635-1639, May 2021.

[29] M. Elhattab, M. A. Arfaoui, C. Assi, and A. Ghrayeb, "Reconfigurable intelligent surface assisted coordinated multipoint in downlink NOMA networks," *IEEE Commun. Lett.*, vol. 25, no. 2, pp. 632-636, Feb. 2021.

[30] H. Wang, S.-H. Leung, and R. Song, "Precoding design for two-cell MIMO-NOMA uplink with CoMP reception," *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2607-2610, Dec. 2018.

[31] H. Wang, C. Liu, Z. Shi, Y. Fu and R. Song, "GSIC for RIS-aided uplink multi-antenna NOMA systems," *IEEE Commun. Lett.*, vol. 26, no. 1, pp. 187-191, Jan. 2022.

[32] P. Marsch and G. P. Fettweis, *Coordinated Multi-Point in Mobile Communications: From Theory to Practice*. Cambridge Univ. Press, 2011.

[33] Z. Shi, H. Wang, Y. Fu, G. Yang, S. Ma, and F. Gao, "Outage analysis of reconfigurable intelligent surface aided MIMO communications with statistical CSI," *IEEE Trans. Wireless Commun.*, early access, pp. 1-17, Jul. 2021.

[34] Y. Cheng, K. H. Li, Y. Liu, K. C. The, and G. K. Karagiannidis, "Non-orthogonal multiple access (NOMA) with multiple intelligent reflecting surfaces," *IEEE Trans. Wireless Commun.*, vol. 20, no. 11, pp. 7184-7195, Nov. 2021.

[35] S. Zhang and R. Zhang, "Intelligent reflecting surface aided multi-user communication: Capacity region and deployment strategy," *IEEE Trans. Commun.*, vol. 69, no. 9, pp. 5790-5806, Sep. 2021.

[36] Q. Wu and R. Zhang, "Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1838-1851, Mar. 2020.

[37] X.-D. Zhang, *Matrix Analysis and Applications*. Cambridge Univ. Press, 2017.

[38] C. Pan, H. Ren, K. Wang, M. Elkashlan, and L. Hanzo, "Intelligent reflecting surface aided MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1719-1734, Aug. 2020.

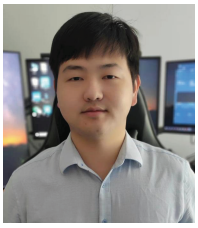
[39] H. Wang, R. Zhang, R. Song, and S.-H. Leung, "A novel power minimization precoding scheme for MIMO-NOMA uplink systems," *IEEE Commun. Lett.*, vol. 22, no. 5, pp. 1106-1109, May 2018.

[40] Z. Wang, L. Liu, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6607-6620, Oct. 2020.

This article has been accepted for publication in IEEE Transactions on Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2022.3177773

Hong Wang received the B.S. degree in communication engineering from Jiangsu University, China, in 2011, and the Ph.D. degree in information and communication engineering from the Nanjing University of Posts and Telecommunications (NUPT), Nanjing, China, in 2016. From 2014 to 2015, he was a Research Assistant with the Department of Electronic Engineering, City University of Hong Kong, Hong Kong. From 2016 to 2018, he was a Senior Research Associate with the State Key Laboratory of Millimeter Waves, Department of Electronic Engineering, City University of Hong Kong. Since 2019, he has been an Associate Professor with the School of Communication and Information Engineering, NUPT. His research interests include broadband wireless communications, particularly in interference management in small cell networks, non-orthogonal multiple access, and intelligent reflecting surface.

Chen Liu (M'18) received the B.E. degree in Electrical and Information Engineering from Nanjing Institute of Technology (Southeast University), China, in 1985 and the M.S. degree in circuits and systems from Anhui University, China, in 1988. He received the Ph.D. degree in signal and information processing from Southeast University, China, in 2005. He joined NUPT in 1988 where he has been a Professor since 2002. His current research interests include massive MIMO systems and reconfigurable intelligent surface.



Zheng Shi received his B.S. degree in communication engineering from Anhui Normal University, China, in 2010 and his M.S. degree in communication and information system from Nanjing University of Posts and Telecommunications (NUPT), China, in 2013. He obtained his Ph.D. degree in Electrical and Computer Engineering from University of Macau, Macao, in 2017. He is currently an Associate Professor with the School of Intelligent Systems Science and Engineering, Jinan University, Zhuhai, China. His current research interests include hybrid automatic repeat request, non-orthogonal multiple access, short-packet communications, intelligent reflecting surface and Internet of Things.

Yaru Fu (S'14-M'18) received her BEng (Hons.) in Telecommunication Engineering from Northeast Electric Power University (NEEPU), MSc (Hons.) in Communication and Information System from Nanjing University of Posts and Telecommunications (NUPT) and Ph.D in Electronic Engineering from City University of Hong Kong (CityU) in 2011, 2014, and 2018, respectively. Then, she joined in the Institute of Network Coding (INC), the Chinese University of Hong Kong (CUHK) as a Post-doctoral Research Assistant. From Sep. 2018 to Jun. 2020, she served as a Research Fellow (Class 2) in Singapore University of Technology and Design. From Feb. 2016 to May. 2016, she was a visiting researcher in Telecom Paris Tech (TPT) and Laboratory of Information, Networking and Communication Sciences (Lincs). She was also an Intern with Nokia Bell Labs, Paris, France. Now, she is an Assistant Professor (first-class) with the School of Science and Technology, Hong Kong Metropolitan University, Hong Kong SAR, China. Her research interests include intelligent wireless communications and networking, caching and recommendation, distributed storage systems, IoT, and URLLC. She is currently serving as an Associate Editor for the IEEE WIRELESS COMMUNICATIONS LETTERS, an Associate Editor for the IEEE NETWORKING LETTERS, and a Review Editor for the FRONTIERS IN COMMUNICATIONS & NETWORKS.

Rongfang Song received the B.S. and M.S. degrees in Telecommunications Engineering and Signals, Circuits and Systems from Nanjing University of Posts and Telecommunications (NUPT), Nanjing, China, in 1984 and 1989, respectively, and the Ph.D. degree in Communications and Information Systems from Southeast University, Nanjing, in 2001. From 2002 to 2003, he was a Research Associate with the Department of Electronic Engineering, City University of Hong Kong, Hong Kong. Since 2002, he has been a Professor with the Department of Telecommunications Engineering, NUPT. His research interests include broadband wireless communications, with current focus on key technologies of B5G/6G.