User Clustering for STAR-RIS Assisted Full-Duplex NOMA Communication Systems

Abdelhamid Salem, *Member, IEEE*, and Kai-Kit Wong, *Fellow, IEEE*,
Chan-Byoung Chae, *Fellow, IEEE*, and Yangyang Zhang

***Abstract*—In contrast to conventional reconfigurable intelligent surface (RIS), simultaneous transmitting and reflecting reconfigurable intelligent surface (STAR-RIS) has been proposed recently to enlarge the serving area from 180^o to 360^o coverage. This work considers the performance of a STAR-RIS aided full-duplex (FD) non-orthogonal multiple access (NOMA) communication systems. The STAR-RIS is implemented at the cell-edge to assist the cell-edge users, while the cell-center users can communicate directly with a FD base station (BS). We first introduce new user clustering schemes for the downlink and uplink transmissions. Then, based on the proposed transmission schemes closed-form expressions of the ergodic rates in the downlink and uplink modes are derived taking into account the system impairments caused by the self interference at the FD-BS and the imperfect successive interference cancellation (SIC). Moreover, an optimization problem to maximize the total sum-rate is formulated and solved by optimizing the amplitudes and the phase-shifts of the STAR-RIS elements and allocating the transmit power efficiently. The performance of the proposed user clustering schemes and the optimal STAR-RIS design are investigated through numerical results.**

***Index Terms*—STAR-RIS, Full-duplex, NOMA, Sum-rate.**

I. Introduction

RECONFIGURABLE intelligent surface (RIS) has been envisioned as a promising technology for sixth-generation (6G) wireless communication networks [1]–[3]. RIS is capable of manipulating the propagation of electromagnetic waves. More specifically, the conventional RIS consists of controllable reflecting elements that can manipulate the phase shifts of the impinging signals to improve the quality of the received signals [1]–[3]. However, the single-faced structure of the conventional RIS limits its service to only half-space, and the base station (BS) and the users should be situated on the same RIS side. To tackle this shortcoming, very recently simultaneous transmitting and reflecting RIS (STAR-RIS) has been proposed and investigated in the literature [4]–[13]. STAR-RIS can simultaneously transmit and reflect the impinging signals to the users located on both sides of the surface, and thus it can extend the serving area from 180^o to 360^o coverage. The concept of STAR-RIS, its potential benefits, and the key differences between conventional RIS and STAR-RIS have been discussed in [4] and [5]. The authors in [4] and [5] proposed three practical STAR-RIS operating protocols, namely, mode switching (MS), time switching (TS), and energy splitting (ES). In MS protocol the STAR-RIS elements are divided into a transmission group and a reflecting group, and in TS protocol the STAR-RIS elements periodically switch between the transmission and reflection modes. Whilst, ES protocol splits the energy of the incident signal on each element into two portions for transmitting and reflecting. In [6] a STAR-RIS-assisted two-user downlink (DL) communication systems has been considered for orthogonal multiple access (OMA) and non-orthogonal multiple access (NOMA) techniques. The performance of STAR-RIS empowered NOMA systems to support low-latency communications has been studied in [7]. Analytical expressions to evaluate the ergodic rate and the coverage probability of a STAR-RIS assisted NOMA multi-cell networks were provided in [8]. In [9] a power minimization problem for STAR-RIS-aided uplink (UL) NOMA systems has been studied. The energy efficiency of a STAR-RIS aided multiple-input and multiple-output (MIMO)-NOMA systems with a user-pairing scheme has been considered in [10]. Approximate analytical expressions of the ergodic rate for a STAR-RIS-aided DL NOMA systems with statistical channel state information (CSI) have been derived in [11]. In [12] the achievable sum rate of a STAR-RIS assisted NOMA systems has been maximized by optimizing the decoding order and the transmission-reflection beamforming. In addition, analytical expressions of the ergodic rate and the outage probability for a pair of users in a STAR-RIS assisted NOMA systems over Rician fading channels have been derived in [13].

Moreover, full-duplex (FD) technique allows wireless devices to simultaneously transmit and receive messages in the same time and frequency resources [14], [15]. Consequently, FD communication systems can offer more flexible usage of the spectrum. Interestingly, the implementation of RIS in FD communication systems has been proposed and studied recently in the literature [16]–[20]. In [16] a RIS has been deployed to cover the dead zones of FD cellular communication systems, the results in this work showed that the spectral efficiency can be doubled compared to the conventional half duplex (HD) systems. In [17] a joint beamforming design for a RIS-aided multiple-antennas FD communication systems has been studied, where the active beamforming at the transmitter and the passive beamforming at the RIS were designed. In [18] the deployment design for a RIS-assisted FD communication system has been

A. Salem is with the Department of Electronic and Electrical Engineering, University College London, WC1E 6BT London, U.K., and also with the Department of Electronic and Electrical Engineering, Benghazi University, Benghazi, Libya (e-mail: a.salem@ucl.ac.uk).

K.-K. Wong is with the department of Electronic and Electrical Engineering, University College London, London, UK, (email: kai-kit.wong@ucl.ac.uk). Kai-Kit Wong is also affiliated with Yonsei Frontier Lab., Yonsei University, Seoul, 03722 Korea.

C.-B. Chae is with the School of Integrated Technology, Yonsei University, Seoul 03722, South Korea, (email: cbchae@yonsei.ac.kr).

Y. Zhang is with the Kuang-Chi Science Ltd., Hong Kong SAR, China.

(Corresponding authors: Kai-Kit Wong and Chan-Byoung Chae).

The work is supported by the Engineering and Physical Sciences Research Council (EPSRC) under grant EP/V052942/1 and by the National Research Foundation of Korea (NRF) Grant through the Ministry of Science and ICT (MSIT), Korea Government, under Grants (RS-2021-0-00486, RS-2024-00428780).

considered, in which an FD transmitter communicates with an UL user and a DL user over the same time and frequency resources through a RIS. A multiple users FD communication system was studied in [19] where a dedicated RIS was assigned to each user in the system. In our previous work [20], we investigated the performance of a STAR-RIS-assisted FD-NOMA pairing communication systems.

Furthermore, it has been shown in the literature that, the detection complexity of NOMA increases as number of the users increases due to the increase in the SIC layers. To overcome this issue, a user clustering scheme has been introduced and investigated [21]–[25]. For instance, in [21] a comprehensive user clustering strategy for DL NOMA systems has been introduced, in which the users grouped in a cluster can receive their signals simultaneously using NOMA technique, and time-division multiple access (TDMA) scheme has been implemented among the different clusters. A novel user clustering scheme for NOMA assisted multi-cell massive MIMO systems was proposed in [22]. Efficient user clustering schemes for UL and DL NOMA transmissions have been designed in [23]. In [24] a user scheduling problem for DL NOMA systems has been studied, in which the BS allocates the spectrum and power resources to a set of users by taking into account the users' fairness. In addition, a user clustering scheme for UL multiple-input single-output (MISO)-NOMA systems has been considered in [25].

Accordingly, this paper considers a STAR-RIS assisted multiple-users FD communication system, where the FD-BS employs NOMA technique to serve the DL and UL users. The cell-edge users communicate with the FD-BS via a passive STAR-RIS, while the cell-center users can communicate directly with the FD-BS. The BS and the STAR-RIS are assumed to know only the statistical channel state information (CSI) and the users' distances. We first present new simple and efficient user-clustering schemes for DL and UL NOMA transmissions. Then, according to the proposed transmission schemes, the ergodic rates for the DL and UL modes are analyzed. In addition, the total sum rates are maximized by optimizing the STAR-RIS elements and the power transmissions. An efficient user clustering and power allocation for both DL and UL users in STAR-RIS assisted FD NOMA systems are the most fundamental design issues. Most of the works in the literature have been focused either on DL or UL scenario, there is no comprehensive investigation to precisely analyze the DL and UL and their respective impact on the user grouping and power allocation problems. In this context, this paper focuses on developing a simple and efficient user clustering and power allocation solutions for STAR-RIS assisted FD NOMA systems. The clustering problem in FD NOMA systems is very hard and challenging , because the clustering should be designed for both DL and UL users simultaneously. The optimal user clustering solution might require an exhaustive search to form a NOMA cluster. It is clear that, the computational complexity of the optimal user clustering is extremely high and impractical with a large number of users. Several methods of user clustering techniques have been considered to efficiently utilize resources and improve the overall sum rate. Matching algorithm can match the users to clusters based on channel conditions. The algorithm uses mathematical methods to determine the best clusters based on a specific criteria. On contrast to these schemes, our proposed algorithm is very simple, efficient and easy to implement comparing with the other schemes that have higher complexity. For clarity, the main contributions are listed as follows:

1) New user clustering schemes for STAR-RIS-aided FD NOMA communication systems are presented and discussed.

2) New closed-form analytical expressions for the ergodic rates of DL and UL users under Rician fading channels are derived. This channel model is more general but also challenging and hard to analyze. The impact of the imperfect SIC and the self interference at the FD-BS are also taken into account in the analysis.

3) The optimal STAR-RIS phase shifts and amplitudes that maximize the ergodic sum-rates are obtained. In addition, sub-optimal designs of the STAR-RIS phase shifts and amplitudes are also provided.

4) An efficient power allocation scheme that enhances the ergodic sum-rates is considered.

5) Monte-Carlo simulations are executed to investigate the performance of the user clustering schemes and the optimal system design.

The results in this work show that increasing the transmit signal to noise ratio (SNR) always improves the achievable rates, and the performance of the cell-edge users can be enhanced by increasing number of the STAR-RIS elements. Furthermore, NOMA user clustering scheme can achieve a higher sum rate than NOMA pairing scheme in the DL mode with perfect and imperfect SIC, when 0% and 10% of the power remains as interference. Whilst, in the UL mode NOMA user clustering scheme outperforms NOMA paring scheme in the perfect SIC case and in the low SNR regime when 10% of the power remains as interference.

II. System Model

Considering a STAR-RIS-assisted multiple-users FD NOMA communication system shown in Fig. 1. The FD-BS is situated at the cell-center with coverage radius, R_t, and the STAR-RIS with N reconfigurable elements is situated at the cell-edge with coverage radius R_r. According to this deployment, the total coverage area R_t can be divided into two areas, the cell-center area with radios R, and the cell-edge area with radios R_r. The FD-BS is equipped with two antennas, one for transmission and one for reception, while the users are equipped with a single antenna[1]. Number of the DL and UL cell-center users are K_{cd} and K_{cu}, respectively, where $K_{cd} + K_{cu} = K_c$, and number of the DL and UL cell-edge users are K_{ed} and K_{eu}, respectively, where $K_{ed} + K_{eu} = K_e$, the total number of the DL and UL users are $K_{cd} + K_{ed} = K_d$ and $K_{cu} + K_{eu} = K_u$, respectively. The users are assumed to be uniformly distributed

[1]Please note that, in this work our main focus is to present a proof of concept for STAR-RIS assisted FD NOMA networks, where our main target is characterizing the corresponding performance and providing impactful and meaningful insights. A single antenna is a tractable scenario that enables us to provide useful insights and gain a better understanding of the properties of STAR-RIS assisted FD NOMA networks. It is worth mentioning that when the BS is equipped with multiple antennas, the resulting multiple-input-single-output (MISO) NOMA system can be decomposed into multiple separate single antenna NOMA systems that are identical to the one considered in this paper. This scenario will be considered in a separate future work, where the derived results in this work can be exploited.

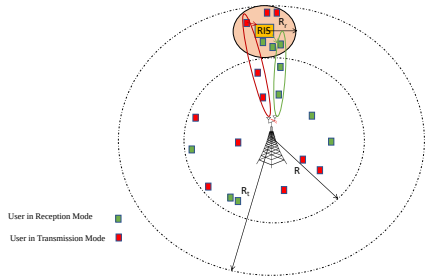


Figure 1: A STAR-RIS assisted FD communication system.

in the areas as shown in Fig 1. The distinguishing between the cell-edge and cell-center can be made based on the received signal strength, SNR, and/or the distance from the base-station. In this work, the users who can communicate directly with the BS and achieve high SINR are classified as cell-center users, while the users who can communicate with the BS through the STAR-RIS and achieve lower SINR are defined as cell-edge users. Rician fading model is assumed for the STAR-RIS related channels, while Rayleigh fading model is assumed for BS to users and user to user channels due to extensive scatterers. It is also assumed that the BS and the RIS can only know the statistical CSI and the users' locations.

In this work, the ES operating protocol is implemented at the STAR-RIS. Thus, the STAR-RIS elements operate in the transmission (t) and reflection (r) modes, and the energy of the incident signal on each element is split into two portions, one for the transmitted signals and one for the reflected signals with the energy-splitting coefficients (amplitude coefficients) ρ_n^r and ρ_n^t ($\rho_n^r + \rho_n^t \leq 1$). The transmission and reflection matrices of the STAR-RIS can be written as $\Theta_k = \text{diag}\left(\rho_1^k\theta_1^k, ..., \rho_N^k\theta_N^k\right), k \in \{t, r\}$, where $\theta_n^k = e^{j\phi_n^k}$, $\rho_n^k \in [0,1]$, $\rho_n^k > 1$ and $\left|\theta_n^k\right| = 1$. By employing STAR-RIS, one side can be assigned to one DL and the other side to the UL mode. With conventional RIS, the RIS should be assigned for UL or DL transmission, this will increase the latency. The selection between RIS and STAR-RIS depends on the specific application requirements, including performance needs, budget constraints, and the acceptable level of complexity for deployment and maintenance.

A. Transmission Model

In the DL transmission, the BS transmits the following signal, $s = \sum_{i=1}^{K_d} \sqrt{\alpha_i} x_{u_{id}}$, where α_i is the power allocation coefficients and $x_{u_{id}}$ is the information signal of the DL user i with unit variance. The received signal at the i^{th} DL cell-center user in can be written as

$$y_{u_{id}} = \sum_{i=1}^{K_d} \sqrt{\alpha_i P_b l_{b,u_{id}}^{-m}} h_{b,u_{id}} x_{u_{id}} + I_{u_{id}} + n_{u_{id}} \quad (1)$$

where P_b is the BS transmit power, $l_{b,u_{id}}^{-m}$ is the path-loss between the BS and the user, m is the path-loss exponent, $h_{b,u_{id}} \sim CN(0,1)$ is the channel between the BS and the user, $n_{u_{id}}$ is the additive white Gaussian noise (AWGN) at the user, $n_{u_{id}} \sim CN\left(0, \sigma_{u_{id}}^2\right)$, and $I_{u_{id}}$ is the interference caused by the UL transmissions and given by

$$I_{u_{id}} = \underbrace{\sum_{j=1}^{K_{cu}} \sqrt{p_{u_{ju}} l_{u_{id},u_{ju}}^{-m}} h_{u_{id},u_{ju}} x_{u_{ju}}}_{\text{UL cell-center users signal}} + \underbrace{\sum_{j=1}^{K_{eu}} \sqrt{p_{u_{ju}} l_{r,u_{ju}}^{-m} l_{r,u_{id}}^{-m}} \mathbf{g}_{r,u_{id}} \Theta_t \mathbf{g}_{r,u_{ju}} x_{u_{ju}}}_{\text{UL cell-edge users signal}} \quad (2)$$

where $p_{u_{iu}}$ and $x_{u_{iu}}$ are the transmit power and the signal of user i, $h_{u_{id},u_{iu}} \sim CN(0,1)$ is the channel between the user and UL cell-center user i, $l_{u_{id},u_{iu}}^{-m}, l_{r,u_{ju}}^{-m} l_{r,u_{id}}^{-m}$ represent the path-loss between the user and j^{th} UL user and the STAR-RIS, and $\mathbf{g}_{r,u_{id}} = \left(\sqrt{\frac{\kappa_{r,u_{id}}}{\kappa_{r,u_{id}}+1}}\bar{\mathbf{g}}_{r,u_{id}} + \sqrt{\frac{1}{\kappa_{r,u_{id}}+1}}\tilde{\mathbf{g}}_{r,u_{id}}\right), \mathbf{g}_{r,u_{ju}} = \left(\sqrt{\frac{\kappa_{r,u_{ju}}}{\kappa_{r,u_{ju}}+1}}\bar{\mathbf{g}}_{r,u_{ju}} + \sqrt{\frac{1}{\kappa_{r,u_{ju}}+1}}\tilde{\mathbf{g}}_{r,u_{ju}}\right)$ are the channel vectors between the RIS and the user, and the RIS and the j^{th}UL cell-edge user, respectively. The received signal at the i^{th} DL cell-edge user can be written as

$$y_{u_{id}} = \sum_{i=1}^{K_d} \sqrt{\alpha_i P_b l_{b,r}^{-m} l_{r,u_{id}}^{-m}} g_{r,u_{id}} \Theta_r g_{b,r} x_{u_{id}} + I_{u_{id}} + n_{u_{id}} \quad (3)$$

where $l_{b,r}^{-m}$ is the path-loss between the BS and the STAR-RIS and $I_{u_{id}}$ is the interference caused by the UL transmissions and given by

$$I_{u_{id}} = \underbrace{\sum_{j=1}^{K_u} \sqrt{p_{u_{ju}} l_{r,u_{id}}^{-m} l_{r,u_{ju}}^{-m}} \mathbf{g}_{r,u_{id}} \Theta_r \mathbf{g}_{r,u_{ju}} x_{u_{ju}}}_{\text{UL users signal}}. \quad (4)$$

In UL mode, the received signal at the BS can be expressed as

$$y_b = \sum_{i=1}^{K_{cu}} \sqrt{p_{u_{iu}} l_{b,u_{iu}}^{-m}} h_{b,u_{iu}} x_{u_{iu}} + \sum_{i=1}^{K_{eu}} \sqrt{p_{u_{iu}} l_{b,r}^{-m} l_{r,u_{iu}}^{-m}} \mathbf{g}_{b,r} \Theta_t \mathbf{g}_{r,u_{iu}} x_{u_{iu}} + I_b + n_b \quad (5)$$

where n_b is the AWGN at the BS, $n_b \sim CN\left(0, \sigma_b^2\right)$ and I_b is the interference term which can be represented by

$$I_b = \underbrace{\sqrt{P_b} h_{b,b} s}_{\text{self interference}} + \underbrace{\sqrt{P_b l_{b,r}^{-m} l_{b,r}^{-m}} \mathbf{g}_{b,r}^H \Theta_t \mathbf{g}_{b,r} s}_{\text{reflection of the DL signal}}. \quad (6)$$

where $h_{b,b}$ is the channel between the transmit and receive antennas. Several self interference suppression (SIS) techniques have been proposed in the literature, such as passive cancellation, analog and digital cancellations, etc [26], [27]. Employing these interference suppression schemes will minimize the self interference to the background noise floor. Following the results in [26], [27], the residual self-interference, $\tilde{s}$, is assumed to follow Gaussian distribution with zero-mean and variance V, $\tilde{s} \sim CN(0, V)$. According to the experimental results, the variance of the residual self interference can be modeled as $V = \beta P_b^{\lambda}$, where β and λ ($0 \leq \lambda \leq 1$) are constants reflect the quality of the cancellation technique.

B. Problem Formulation

In the considered model, the FD-BS employs NOMA scheme to serve the DL and UL users. To reduce the detection complexity of NOMA technique, the DL and UL users are divided into orthogonal clusters. Number of the DL clusters, $1 \leq M_d \leq \frac{K_d}{2}$, and UL clusters, $1 \leq M_u \leq \frac{K_u}{2}$. In addition, number of the users in each DL cluster and UL cluster are denoted by κ_d and κ_u, respectively, where $2 \leq \kappa_d \leq K_d$ and $2 \leq \kappa_u \leq K_u$. The joint user-clustering, STAR-RIS design, and power allocation to maximize the total DL and UL sum rates can be formulated as

$$\max_{\rho,\theta,\mathbf{p},\beta} \sum_{j=1}^{\frac{K_d}{2}} \sum_{i=1}^{K_d} \beta_{d_{i,j}} \bar{R}_{u_{i,jd}} + \sum_{j=1}^{\frac{K_u}{2}} \sum_{i=1}^{K_u} \beta_{u_{i,j}} \bar{R}_{u_{i,ju}}$$

$$(C.1) \quad \text{s.t} \sum_{j=1}^{\frac{K_d}{2}} \beta_{d_{i,j}} \bar{R}_{u_{i,jd}} > \hat{R}_{d_i}, \sum_{j=1}^{\frac{K_u}{2}} \beta_{u_{i,j}} \bar{R}_{u_{i,ju}} > \hat{R}_{u_i}$$

$$(C.2) \quad R_{u_{kd} \to u_{id}} \geqslant R_{u_{id}}, k > i$$

$$(C.3) \quad \beta_{d_{i,j}} \in \{0,1\}, \beta_{u_{i,j}} \in \{0,1\}, \forall i,j$$

$$(C.4) \quad \sum_{j=1}^{\frac{K_d}{2}} \beta_{d_{i,j}} = 1, \sum_{j=1}^{\frac{K_u}{2}} \beta_{u_{i,j}} = 1, \forall i$$

$$(C.5) \quad 2 \leq \sum_{i=1}^{K_d} \beta_{d_{i,j}} \leq K_d, 2 \leq \sum_{i=1}^{K_u} \beta_{u_{i,j}} \leq K_u, \forall j$$

$$(C.6) \quad (\rho_n^r) + (\rho_n^t) = 1, \forall n \in N$$

$$(C.7) \quad \rho_n^k \geq 0, \left|\theta_n^k\right| = 1, \forall n \in N$$

$$(C.8) \quad \sum_{i=1}^{K_d} P_{b_i} \leq P_b, P_{b_i} > 0, 0 < p_{u_{iu}} \leq p_{u_m} \qquad (7)$$

where $\bar{R}_{u_{i,jd}}$ is the average rate of the i^{th} DL user in cluster j, $\bar{R}_{u_{i,ju}}$ is the average rate of the i^{th} UL user in cluster j, and

$$\beta_{p_{i,j}} = \begin{cases} 1, & \text{if a user } i \text{ is grouped into cluster } j \\ 0, & \text{otherwise} \end{cases}$$

$\mathbf{p} = \left[p_{u_1}, ..., p_{u_{K_u}}, P_{b_1}, ..., P_{b_{K_d}}\right]^T$, $P_{b_i} = \alpha_i P_b$, P_b and p_{u_m} are the maximum transmission power at the BS and the user. The first constraint $(C.1)$ provides the data rate requirements of the DL and UL users, and the second constraint $(C.2)$ provides the SIC implementation condition, while the constraint $(C.3)$ explains that $\beta_{d_{i,j}}$ and $\beta_{u_{i,j}}$ are integer variables. The constraints $(C.4)$ and $(C.5)$ are required to ensure that each user is assigned to only one cluster and that each cluster has at least two users, constraints $(C.6)$ and $(C.7)$ for the amplitude and phase shift on each STAR-RIS element. In addition, constraint $(C.8)$presents the power constraint of the BS and the users.

As we can observe from (7), it is extremely difficult to find the solution of the problem due to its non-convexity in nature. Specifically, an exhaustive search algorithm is required to find the optimal solution for NOMA user clustering scheme. Thus, for each DL and UL user all possible combinations should be considered. For instance, in the DL number of the possible combinations is: $\sum_{i=1}^{K_d} \begin{pmatrix} K_d \\ i \end{pmatrix}$. It is clear that, the computational complexity of the optimal user clustering is extremely high and impractical with a large number of users. In addition, considering the two variables, ρ, θ, to obtain the optimal STAR-RIS design makes the problem more complicated. In the following sections we divide the main problem into two sub-problems. We first develop a less complex and efficient solution for user clustering scheme, and then we consider the optimal system design.

III. User Clustering Scheme

A user clustering scheme is an efficient solution to reduce the decoding and implementation complexities of SIC in NOMA systems. In the user clustering scheme number of the signals decoded by SIC can be reduced by distributing the users into small groups [28], [29]. In this section, we introduce low-complexity sub-optimal user clustering schemes for a STAR-RIS-aided FD-NOMA systems. The proposed scheme exploits the users locations, and the distances between the users, STAR-RIS and the FD-BS.

A. Key Principles

i) DL: Following the principles of NOMA, the DL users are sorted in descending order based on their distances to the BS, $l_{b,u_{Kd}} > .. > l_{b,u_{1d}}$ where $l_{b,u_{kd}}$is the distance between the BS and user k[2]. After applying SIC, the rate of the closest users to the BS (the strong/cell-center users) will depend essentially on their channel variances and the amount of power allocated to each user. Thus, it is reasonable and beneficial to distribute the strong users into different clusters to enhance the achievable sum-rate. For the far users (the weak/cell-edge users), it is more reasonable and useful to include them with the strong users. This is because the strong users can achieve high rates with low power portions, and thus large portion of power can be allocated to the weak users. Accordingly, the key idea here is to group the strongest users and the weakest users into the same cluster [22], [33].

ii) UL: Similarly, in UL the users are sorted in descending order according to their distances to the BS, $l_{b,u_{Ku}} > ... > l_{b,u_{1u}}$. In UL NOMA, the strong users have less impact on the performance of the weak users, because the strongest users' messages will be canceled/reduced using SIC scheme. Thus, the users can transmit their messages using the highest available power. Accordingly, it is useful to cluster the strong users together to enhance the achievable sum-rate of the cluster [22], [33].

B. User Clustering Algorithm

i) DL: To simplify NOMA operation, the DL cell-center users are virtually partitioned into two groups, group 1 ($\mathcal{G}_1^d$) and group 2 ($\mathcal{G}_2^d$), their locations are as much different as possible [33], [34]. The users in $\mathcal{G}_1^d$ are the nearest to the BS and the users in $\mathcal{G}_2^d$ are the farthest users to the BS. Number of the users in group $\mathcal{G}_1^d$ is K_{d_1}, while number of the users in group $\mathcal{G}_2^d$ is K_{d_2}. This partitioning of the users results in three different groups in total, $\mathcal{G}_1^d$, $\mathcal{G}_2^d$ and the cell-edge users group $\mathcal{G}_3^d$. Following the key principles above, we sort the DL users in each group in descending order, $l_{b,u_{K_{di}}} > .. > l_{b,u_{1di}}$, where $i \in \{1,2,3\}$ is

[2]NOMA clustering scheme proposed in this work does not required perfect knowledge of instantaneous CSI at the BS. It is assumed that only statistical CSI is known, thus the decoding/users order is determined based on their average channel gains. e.g, path loss. Generally, the path loss remains stable over time periods compared to small-scale fading. The practice of using distance-based ordering in NOMA systems is extensively investigated in [30]–[32]. This scheme can also simplify the power-allocation design for NOMA [9].

Algorithm 1 User Clustering Algorithm.

1. Calculate the distances between the users and the BS.
2. Sort the DL and UL users according to their distances to the BS.
3. Divide the users into groups, and assign the users in each group: $\underbrace{l_{b,u_{1d1}}^{-m} > l_{b,u_{2d1}}^{-m} > ..}_{\text{DL users in } \mathcal{G}_1} > \underbrace{l_{b,u_{1d2}}^{-m} > l_{b,u_{2d2}}^{-m} > ..}_{\text{DL users in } \mathcal{G}_2}$ $\underbrace{.. > l_{b,u_{K_{ed}-1}}^{-m} > l_{b,u_{K_{ed}}}^{-m}}_{\text{DL users in } \mathcal{G}_3}$ and $\underbrace{l_{b,u_{1u1}}^{-m} > l_{b,u_{2u1}}^{-m} > ..}_{\text{UL users in } \mathcal{G}_1} > \underbrace{l_{b,u_{1u2}}^{-m} > l_{b,u_{2u2}}^{-m} > .}_{\text{UL users in } \mathcal{G}_2}$ $\underbrace{.. > l_{b,u_{K_{eu}-1}}^{-m} > l_{b,u_{K_{eu}}}^{-m}}_{\text{UL users in } \mathcal{G}_3}$.
   4.1 Group the DL users into clusters: 1st cluster $= \left\{l_{b,u_{1d1}}^{-m}, l_{b,u_{1d2}}^{-m}, l_{b,u_{K_{ed}}}^{-m}\right\}$, .. and last cluster $= \left\{l_{b,u_{K_{d1}}}^{-m}, l_{b,u_{K_{d2}}}^{-m}, l_{b,u_{K_{d1}+K_{d2}+1}}^{-m}\right\}$.
   4.2 Group the UL users into clusters:1st cluster $= \left\{l_{b,u_{1u1}}^{-m}, l_{b,u_{1u2}}^{-m}, l_{b,u_{1u3}}^{-m}\right\}$, .. and last cluster $= \left\{l_{b,u_{K_{u1}}}^{-m}, l_{b,u_{K_{u2}}}^{-m}, l_{b,u_{K_{eu}}}^{-m}\right\}$.

the group index. Accordingly, the strongest users in $\mathcal{G}_1^d$ and $\mathcal{G}_2^d$ and the weakest user in $\mathcal{G}_3^d$ are grouped into the same cluster, while the second strongest users in $\mathcal{G}_1^d$ and $\mathcal{G}_2^d$ and the second weakest user in $\mathcal{G}_3^d$ are grouped into another cluster, and so on[3].

ii) UL: Similarly, the UL cell-center users are virtually partitioned into two groups, group 1 ($\mathcal{G}_1^u$) and group 2 ($\mathcal{G}_2^u$), their locations are as much different as possible [33], [34]. Number of the users in group $\mathcal{G}_1^u$ is K_{u_1} while number of the users in group $\mathcal{G}_2^u$ is K_{u_2}. This results in three different groups in total, $\mathcal{G}_1^u$, $\mathcal{G}_2^u$ and the cell-edge users group $\mathcal{G}_3^u$. According to the key principles above, we sort the UL users in each group in descending order, $l_{b,u_{K_{ui}}} > ... > l_{b,u_{1ui}}$, where $i \in \{1,2,3\}$. Thus, the strongest users in each group are included in a cluster, the second strongest users in each group are grouped in another cluster, and so on.

All steps of the proposed algorithm are summarized in Algorithm 1. In addition, for clarity in Fig. 2, we illustrate the user clustering scheme for 6 cell-center users, and 3 cell-edge users in DL and UL modes[4].

IV. Ergodic Rate Analysis

In this section, based on the proposed user clustering scheme we derive the ergodic rates of the DL and UL users, which will be used later to find the optimal system design.

A. DL

The BS transmits the following signal to a cluster in the system, $s = \sum_{i=1}^{3} \sqrt{\alpha_i} x_{u_{id}}$, where α_i is the power allocation coefficients with $\alpha_1 + \alpha_2 + \alpha_3 \leq 1$,$\alpha_1 < \alpha_2 < \alpha_3$.

[3]Note that the analysis in this work is also applicable when the DL and UL groups have different numbers of users.

[4]Increasing number of the users in a cluster increases the detection complexity but also reduces the usage of the resources/time slots. Thus, moving a step above NOMA paring scheme will provide better usage of the resources, which can compensate for the small errors in the SIC stage.

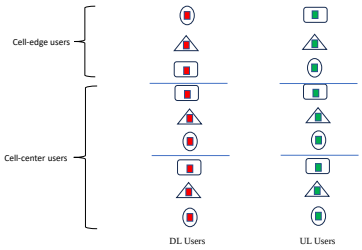


Figure 2: Illustration of user clustering scheme to serve 6 cell-center users and 3 cell-edge users in DL and UL modes.

1) Cell-center, Group $\mathcal{G}_1$: The SINR at the DL strongest user in a cluster, user 1, can be written as

$$\gamma_{u_{1d}} = \frac{P_{b_1} A_{u_{1d}}}{\Xi P_b \left(\alpha_2 + \alpha_3\right) A_{u_{1d}} + B_{u_{1d}} + C_{u_{1d}} + D_{u_{1d}} + \sigma_{u_{1d}}^2} \tag{8}$$

where

$A_{u_{1d}} = \left|\sqrt{l_{b,u_{1d}}^{-m}} h_{b,u_{1d}}\right|^2$, $B_{u_{1d}} = p_{u_{1u}} \left|\sqrt{l_{u_{1d},u_{1u}}^{-m}} h_{u_{1d},u_{1u}}\right|^2$,

$C_{u_{1d}} = p_{u_{2u}} \left|\sqrt{l_{u_{1d},u_{2u}}^{-m}} h_{u_{1d},u_{2u}}\right|^2$,

$D_{u_{1d}} = p_{u_{3u}} l_{r,u_{3u}}^{-m} l_{r,u_{1d}}^{-m} \left|\mathbf{g}_{r,u_{1d}} \Theta_t \mathbf{g}_{r,u_{3u}}\right|^2$,

and $P_{b_1} = \alpha_1 P_b$ and $0 \leq \Xi \leq 1$ is the SIC error factor which represents a fraction of the power that remains as interference due to imperfect SIC.

Theorem 1. *The ergodic rate of the DL strongest user in a cluster can be evaluated by*

$$\mathcal{E}\left[R_{u_{1d}}\right] \approx \frac{1}{M_d} \log_2 \left(1 + \frac{P_{b_1} x_{1_{u_{1d}}}}{\Xi P_b \left(\alpha_2 + \alpha_3\right) x_{1_{u_{1d}}} + p_{u_{1,2}} y_{1_{u_{1d}}} + p_{u_{3u}} \omega_{r,u_{3u}}^{r,u_{1d}} y_{2_{u_{1d}}} + \sigma_{u_{1d}}^2}\right) \tag{9}$$

where $x_{1_{u_{1d}}}, y_{1_{u_{1d}}}$, and $y_{2_{u_{1d}}}$ are defined in (10)-(12), k_{cd_1} is the user order $1 \leq k_{cd_1} \leq K_{d_1}$, $\mathcal{H}(.)$ is the hypergeometric function, $p_{u_{1,2}} = (p_{u_{1u}} + p_{u_{2u}})$, $a_{1_{u1d}} = \frac{\sqrt{\pi}\Gamma\left(1-k_{cd_1}+K_{cd}\right)2K_{cd}!}{4\left(k_{cd_1}-1\right)!\left(K_{cd}-k_{cd_1}\right)!}$, $a_{1_{u3u}} = \frac{\sqrt{\pi}\Gamma\left(1-k_{eu_3}+K_{eu}\right)2K_{eu}!}{4\left(k_{eu_3}-1\right)!\left(K_{eu}-k_{eu_3}\right)!}$, $\omega_{r,u_{3u}}^{r,u_{1d}} = \left(\varpi_{r,u_{3u}}^{r,u_{1d}} \xi_1 + \hat{\varpi}_{r,u_{3u}}^{r,u_{1d}}\right)$, $\xi_1 = \left|\bar{\mathbf{g}}_{r,u1d} \Theta_t \bar{\mathbf{g}}_{r,u_{3u}}\right|^2$, $\varpi_y^x = \frac{\kappa_x}{\kappa_x+1} \frac{\kappa_y}{\kappa_y+1}$, $\hat{\varpi}_y^x = \frac{\kappa_x}{\kappa_x+1} \frac{\sum_{n=1}^{N} \left|\rho_n^k\right|^2}{\kappa_y+1} + \frac{\kappa_y}{\kappa_y+1} \frac{\sum_{n=1}^{N} \left|\rho_n^k\right|^2}{\kappa_x+1} + \frac{1}{\kappa_x+1} \frac{\sum_{n=1}^{N} \left|\rho_n^k\right|^2}{\kappa_y+1}$.

Proof: The proof is provided in Appendix A. ■

2) Cell-center, Group $\mathcal{G}_2$: The SINR at the user can be written as

$$\gamma_{u_{2d}} = \frac{\alpha_2 P_b A_{u_{2d}}}{P_b A_{u_{2d}} \left(\Xi \alpha_3 + \alpha_1\right) + B_{u_{2d}} + C_{u_{2d}} + D_{u_{2d}} + \sigma_{u_{2d}}^2} \tag{13}$$

where

$A_{u_{2d}} = \left|\sqrt{l_{b,u_{2d}}^{-m}} h_{b,u_{2d}}\right|^2$, $B_{u_{2d}} = p_{u_{1u}} \left|\sqrt{l_{u_{2d},u_{1u}}^{-m}} h_{u_{2d},u_{1u}}\right|^2$,

$$x_{1_{u_{1d}}} = 2a_{1_{u1d}}\Gamma\left(k_{cd_1}\right)\mathcal{H}\left(\left\{k_{cd_1}, \frac{1+m}{2}, \frac{m}{2}\right\}, \left\{\frac{1}{2}, 1+K_{cd}\right\}, R^2\right)$$
$$-mRa_{1_{u1d}}\Gamma\left(\frac{1}{2}+k_{cd_1}\right)\mathcal{H}\left(\left\{\frac{1}{2}+k_{cd_1}, \frac{1+m}{2}, \frac{2+m}{2}\right\}, \left\{\frac{3}{2}, \frac{3}{2}+K_{cd}\right\}, R^2\right) \tag{10}$$

$$y_{1_{u_{1d}}} = \frac{2}{\left(2-3m+m^2\right)R^2} - \frac{2\mathcal{H}\left(\left\{\frac{1}{2}, -1+\frac{m}{2}, -\frac{1}{2}+\frac{m}{2}\right\}, \left\{\frac{-1}{2}, 1\right\}, 4R^2\right)}{\left(2-3m+m^2\right)R^2} - \mathcal{H}\left(\left\{\frac{3}{2}, \frac{1}{2}+\frac{m}{2}, \frac{m}{2}\right\}, \left\{\frac{1}{2}, 3\right\}, 4R^2\right)$$
$$+\frac{64mR\,\mathcal{H}\left(\left\{2, \frac{1}{2}+\frac{m}{2}, 1+\frac{m}{2}\right\}, \left\{\frac{3}{2}, \frac{7}{2}\right\}, 4R^2\right)}{15\pi} - \frac{64mR\,\mathcal{H}\left(\left\{2, \frac{1}{2}+\frac{m}{2}, 1+\frac{m}{2}\right\}, \left\{\frac{5}{2}, \frac{5}{2}\right\}, 4R^2\right)}{9\pi} \tag{11}$$

$$y_{2_{u_{1d}}} = \left(2a_{1_{u3u}}\Gamma\left(k_{eu_3}\right)\mathcal{H}\left(\left\{k_{eu_3}, \frac{1+m}{2}, \frac{m}{2}\right\}, \left\{\frac{1}{2}, 1+K_{eu}\right\}, R_r^2\right)\right.$$
$$\left.-mR_ra_{1_{u3u}}\Gamma\left(\frac{1}{2}+k_{eu_3}\right)\mathcal{H}\left(\left\{\frac{1}{2}+k_{eu_3}, \frac{1+m}{2}, \frac{2+m}{2}\right\}, \left\{\frac{3}{2}, \frac{3}{2}+K_{eu}\right\}, R_r^2\right)\right)$$
$$\times\sum_{j=1}^{C}\mathrm{H}_j\,\left(1+\left(R\,r_j+R\right)\right)^{-m}\frac{2\left(R\,r_j+R\right)}{\pi R^2}\cos^{-1}\left(\frac{1}{\left(R\,r_j+R\right)}\left(r_1+\frac{\left(\left(R\,r_j+R\right)^2-r_1^2\right)}{2\left(R+r_1\right)}\right)\right) \tag{12}$$

$$x_{1_{u_{2d}}} = 2a_{1_{u2d}}\Gamma\left(k_{cd_2}\right)\mathcal{H}\left(\left\{k_{cd_2}, \frac{1+m}{2}, \frac{m}{2}\right\}, \left\{\frac{1}{2}, 1+K_{cd}\right\}, R^2\right)$$
$$-mRa_{1_{u2d}}\Gamma\left(\frac{1}{2}+k_{cd_2}\right)\mathcal{H}\left(\left\{\frac{1}{2}+k_{cd_2}, \frac{1+m}{2}, \frac{2+m}{2}\right\}, \left\{\frac{3}{2}, \frac{3}{2}+K_{cd}\right\}, R^2\right) \tag{15}$$

$C_{u_{2d}} = p_{u_{2u}}\left|\sqrt{l_{u_{2d},u_{2u}}^{-m}}h_{u_{2d},u_{2u}}\right|^2,$
$D_{u_{2d}} = p_{u_{3u}}l_{r,u3u}^{-m}l_{r,u2d}^{-m}\left|\mathbf{g}_{r,u_{2d}}\Theta_t\mathbf{g}_{r,u_{3u}}\right|^2.$

Theorem 2. *The ergodic rate of the DL second strongest user in a cluster can be evaluated by*

$$\mathcal{E}\left[R_{u_{2d}}\right] \approx \frac{1}{M_d}\log_2\left(1+\frac{P_{b_2}x_{1_{u_{2d}}}}{P_bx_{1_{u_{2d}}}\left(\Xi\alpha_3+\alpha_1\right)+p_{u_{1,2}}y_{1_{u_{2d}}}+p_{u_{3u}}\omega_{r,u_{3u}}^{r,u_{2d}}y_{2_{u_{2d}}}+\sigma_{u_{2d}}^2}\right) \tag{14}$$

where $x_{1_{u_{2d}}}$ is defined in (15), $y_{1_{u_{2d}}} = y_{1_{u_{1d}}}$, $y_{2_{u_{2d}}} = y_{2_{u_{1d}}}$, k_{cd_2} is the user order $K_{d_1}+1 \le k_{cd_2} \le K_{cd}$, $\omega_{r,u_{3u}}^{r,u_{2d}} = \left(\varpi_{r,u_{3u}}^{r,u_{2d}}\xi_2+\hat{\varpi}_{r,u_{3u}}^{r,u_{2d}}\right)$, $\xi_2 = \left|\bar{\mathbf{g}}_{r,u2d}\Theta_t\bar{\mathbf{g}}_{r,u_{3u}}\right|^2$ and $a_{1_{u2d}} = \frac{\sqrt{\pi}\Gamma\left(1-k_{cd_2}+K_{cd}\right)2K_{cd}!}{4\left(k_{cd_2}-1\right)!\left(K_{cd}-k\right)!}$.

Proof: The proof is provided in Appendix B. ∎

3) Cell-edge, Group $\mathcal{G}_3$: The SINR at the user can be written as

$$\gamma_{u_{3d}} = \frac{A_{u_{3d}}}{P_{b_{1,2}}A_{u_{3d}}+B_{u_{3d}}+C_{u_{3d}}+D_{u_{3d}}+\sigma_{u_{3d}}^2} \tag{16}$$

where $A_{u_{3d}} = \alpha_3P_bl_{b,r}^{-m}l_{r,u_{3d}}^{-m}\left|\mathbf{g}_{r,u_{3d}}\Theta_r\mathbf{g}_{b,r}\right|^2$, $P_{b_{1,2}} = \left(\alpha_1+\alpha_2\right)P_b$,
$B_{u_{3d}} = p_{u_{1u}}l_{r,u_{3d}}^{-m}l_{r,u_{1u}}^{-m}\left|\mathbf{g}_{r,u_{3d}}\Theta_r\mathbf{g}_{r,u_{1u}}\right|^2,$
$C_{u_{3d}} = p_{u_{2u}}l_{r,u_{3d}}^{-m}l_{r,u_{2u}}^{-m}\left|\mathbf{g}_{r,u_{3d}}\Theta_r\mathbf{g}_{r,u_{2u}}\right|^2,$
$D_{u_{3d}} = p_{u_{3u}}l_{r,u_{3d}}^{-m}l_{r,u_{3u}}^{-m}\left|\mathbf{g}_{r,u_{3d}}\Theta_r\mathbf{g}_{r,u_{3u}}\right|^2.$

Theorem 3. *The ergodic rate of the DL weakest user in a cluster can be evaluated by*

$$\mathcal{E}\left[R_{u_{3d}}\right] \approx \frac{1}{M_d}\log_2\left(1+\frac{P_{b_3}x_{u_{3d}}}{P_{b_{1,2}}x_{u_{3d}}+b_1y_{1_{u_{3d}}}+b_2y_{2_{u_{3d}}}+\sigma_{u_{3d}}^2}\right) \tag{17}$$

where $x_{u_{3d}} = l_{b,r}^{-m}\omega_{b,r}^{r,u_{3d}}x_{1_{u_{3d}}}$, $b_1 = p_{u_{1u}}\omega_{r,u_{1u}}^{r,u_{3d}}+p_{u_{2u}}\omega_{r,u_{2u}}^{r,u_{3d}}$, $b_2 = p_{u_{3u}}\omega_{r,u_{3u}}^{r,u_{3d}}$, $x_{1_{u_{3d}}}, y_{1_{u_{3d}}}, y_{2_{u_{3d}}}$ are defined in (18)-(20), k_{ed_3} is the user order $1 \le k_{ed_3} \le K_{ed}$, $\omega_{b,r}^{r,u_{3d}} = \varpi_{b,r}^{r,u_{3d}}\xi_3+\hat{\varpi}_{b,r}^{r,u_{3d}}$ $\omega_{r,u_{1u}}^{r,u_{3d}} = \varpi_{r,u_{1u}}^{r,u_{3d}}\xi_4+\hat{\varpi}_{r,u_{1u}}^{r,u_{3d}}$, $\omega_{r,u_{2u}}^{r,u_{3d}} = \varpi_{r,u_{2u}}^{r,u_{3d}}\xi_5+\hat{\varpi}_{r,u_{2u}}^{r,u_{3d}}$ $\omega_{r,u_{3u}}^{r,u_{3d}} = \varpi_{r,u_{3u}}^{r,u_{3d}}\xi_6+\hat{\varpi}_{r,u_{3u}}^{r,u_{3d}}$ $\xi_3 = \left|\bar{\mathbf{g}}_{r,u3d}\Theta_r\bar{\mathbf{g}}_{b,r}\right|^2$, $\xi_4 = \left|\bar{\mathbf{g}}_{r,u3d}\Theta\bar{\mathbf{g}}_{r,u_{1u}}\right|^2$, $\xi_5 = \left|\bar{\mathbf{g}}_{r,u3d}\Theta\bar{\mathbf{g}}_{r,u_{2u}}\right|^2$, $\xi_6 = \left|\bar{\mathbf{g}}_{r,u3d}\Theta\bar{\mathbf{g}}_{r,u_{3u}}\right|^2$, $a_{1_{u3d}} = \frac{\sqrt{\pi}\Gamma\left(1-k_{ed_3}+K_{ed}\right)2K_{ed}!}{4\left(k_{ed_3}-1\right)!\left(K_{ed}-k_{ed_3}\right)!}$, and $a_{2_{u3d}} = \frac{\sqrt{\pi}\Gamma\left(1-k_{eu_1}+K_{eu}\right)2K_{eu}!}{4\left(k_{eu_1}-1\right)!\left(K_{eu}-k_{eu_1}\right)!}$.

Proof: The proof is provided in Appendix C. ∎

B. UL

Based on the proposed UL transmission scheme, and (5), we can derive the ergodic rates of an UL cluster as follows.

1) Cell-center, Group $\mathcal{G}_1$: The SINR to detect the signal of the UL strongest user in a cluster, user 1, can be written as,

$$\gamma_{u_{1u}} = \frac{p_{u_{1u}}A_{u_{1u}}}{p_{u_{2u}}B_{u_{1u}}+p_{u_{3u}}C_{u_{1u}}+P_bD_{u_{1u}}+V+\sigma_{u_{1u}}^2} \tag{21}$$

$$x_{1_{u_{3d}}} = 2a_{1_{u3d}}\Gamma\left(k_{ed_3}\right)\mathcal{H}\left(\left\{k_{ed_3}, \frac{1+m}{2}, \frac{m}{2}\right\}, \left\{\frac{1}{2}, 1+K_{ed}\right\}, R_r^2\right)$$
$$-mR_r a_{1_{u3d}}\Gamma\left(\frac{1}{2}+k_{ed_3}\right)\mathcal{H}\left(\left\{\frac{1}{2}+k_{ed_3}, \frac{1+m}{2}, \frac{2+m}{2}\right\}, \left\{\frac{3}{2}, \frac{3}{2}+K_{ed}\right\}, R_r^2\right) \tag{18}$$

$$y_{1_{u_{3d}}} = x_{1_{u_{3d}}}\sum_{j=1}^{C}\mathrm{H}_j\ \left(1+\left(R\,r_j+R\right)\right)^{-m}\frac{2\left(R\,r_j+R\right)}{\pi R^2}\cos^{-1}\left(\frac{1}{\left(R\,r_j+R\right)}\left(r_1+\frac{\left(\left(R\,r_j+R\right)^2-r_1^2\right)}{2\left(R+r_1\right)}\right)\right) \tag{19}$$

$$y_{2_{u_{3d}}} = x_{1_{u_{3d}}}a_{2_{u3d}}2\Gamma\left(k_{eu_1}\right)\mathcal{H}\left(\left\{k_{eu_1}, \frac{1+m}{2}, \frac{m}{2}\right\}, \left\{\frac{1}{2}, 1+K_{eu}\right\}, R_r^2\right)$$
$$-mR_r x_{1_{u_{3d}}}a_{2_{u3d}}\Gamma\left(\frac{1}{2}+k_{eu_1}\right)\mathcal{H}\left(\left\{\frac{1}{2}+k_{eu_1}, \frac{1+m}{2}, \frac{2+m}{2}\right\}, \left\{\frac{3}{2}, \frac{3}{2}+K_{eu}\right\}, R_r^2\right) \tag{20}$$

$$\chi_{u_{iu}} = 2a_{i_{u1u}}\Gamma\left(k_{cu_i}\right)\mathcal{H}\left(\left\{k_{cu_i}, \frac{1+m}{2}, \frac{m}{2}\right\}, \left\{\frac{1}{2}, 1+K_{cu}\right\}, R^2\right)$$
$$-mRa_{i_{u1u}}\Gamma\left(\frac{1}{2}+k_{cu_i}\right)\mathcal{H}\left(\left\{\frac{1}{2}+k_{cu_i}, \frac{1+m}{2}, \frac{2+m}{2}\right\}, \left\{\frac{3}{2}, \frac{3}{2}+K_{cu}\right\}, R^2\right) \tag{23}$$

$$y_{2_{u_{1u}}} = 2a_{3_{u1u}}\Gamma\left(k_{eu_1}\right)\mathcal{H}\left(\left\{k_{eu_1}, \frac{1+m}{2}, \frac{m}{2}\right\}, \left\{\frac{1}{2}, 1+K_{eu}\right\}, R_r^2\right)$$
$$-mR_r a_{3_{u1u}}\Gamma\left(\frac{1}{2}+k_{eu_1}\right)\mathcal{H}\left(\left\{\frac{1}{2}+k_{eu_1}, \frac{1+m}{2}, \frac{2+m}{2}\right\}, \left\{\frac{3}{2}, \frac{3}{2}+K_{eu}\right\}, R_r^2\right) \tag{24}$$

$$y_{3_{u_{1u}}} = l_{b,r}^{-m}l_{b,r}^{-m}\left(\left(\frac{\kappa_{b,r}}{\kappa_{b,r}+1}\right)^2\xi_8 + 2\frac{\kappa_{b,r}}{\kappa_{b,r}+1}\frac{1}{\kappa_{b,r}+1}\sum_{n=1}^{N}\left|\rho_n^k\right|^2 + \left(\frac{1}{\kappa_{b,r}+1}\right)^2\right.$$
$$\left.\left(2\sum_{n=1}^{N}\left|\rho_n^k\right|^2 + \sum_{n_1=1}^{N}\sum_{n_2\neq n_1}^{N}\left(\rho_{n_1}^k e^{j\phi_{n_1}^k}\right)\left(\rho_{n_2}^k e^{j\phi_{n_2}^k}\right)^H\right) + 2\frac{\kappa_{b,r}}{\kappa_{b,r}+1}\frac{1}{\kappa_{b,r}+1}\left(\hat{\zeta}\sum_{n=1}^{N}\left(\rho_n^k e^{j\phi_n^k}\right)^H\right)\right) \tag{25}$$

where $A_{u_{1u}} = \left|\sqrt{l_{b,u_{1u}}^{-m}}h_{b,u_{1u}}\right|^2$, $B_{u_{1u}} = \left|\sqrt{l_{b,u_{2u}}^{-m}}h_{b,u_{2u}}\right|^2$ $C_{u_{1u}} = l_{b,r}^{-m}l_{r,u_{3u}}^{-m}\left|\mathbf{g}_{b,r}\Theta_t\mathbf{g}_{r,u_{3u}}\right|^2$, $D_{u_{1u}} = l_{b,r}^{-m}l_{b,r}^{-m}\left|\mathbf{g}_{b,r}^H\Theta_t\mathbf{g}_{b,r}\right|^2$.

Theorem 4. *The ergodic rate of the UL strongest user in a cluster can be evaluated by*

$$\mathcal{E}\left[R_{u_{1u}}\right] \approx \frac{1}{M_u}\log_2\left(1+\frac{p_{u_{1u}}\chi_{u_{1u}}}{p_{u_{2u}}\chi_{u_{2u}}+p_{u_{3u}}\omega_{b,r}^{r,u_{3u}}y_{2_{u_{1u}}}+P_b y_{3_{u_{1u}}}+V+\sigma_b^2}\right) \tag{22}$$

where $\chi_{u_{iu}}, y_{2_{u_{1u}}}$ and $y_{3_{u_{1u}}}$ are defined in (23)-(25), $k_{cu_1}$is the strongest cell-center user order $1 \le k_{cu_1} \le K_{u_1}$, $k_{cu_2}$is the second strongest cell-center user order $K_{u_1}+1 \le k_{cu_2} \le K_{cu}$, k_{eu_1} is the UL cell-edge user order $1 \le k_{eu_1} \le K_{eu}$, $\omega_{b,r}^{r,u_{3u}} = \varpi_{b,r}^{r,u_{3u}}\xi_7 + \hat{\varpi}_{b,r}^{r,u_{3u}}$, $a_{1_{u1u}} = \frac{\sqrt{\pi}\Gamma\left(1-k_{cu_1}+K_{cu}\right)2K_{cu}!}{4\left(k_{cu_1}-1\right)!\left(K_{cu}-k\right)!}$, $a_{2_{u1u}} = \frac{\sqrt{\pi}\Gamma\left(1-k_{cu_2}+K_{cu}\right)2K_{cu}!}{4\left(k_{cu_2}-1\right)!\left(K_{cu}-k_{cu_2}\right)!}$, $a_{3_{u1u}} = \frac{\sqrt{\pi}\Gamma\left(1-k_{eu_1}+K_{eu}\right)2K_{eu}!}{4\left(k_{eu_1}-1\right)!\left(K_{eu}-k_{eu_1}\right)!}$, $\xi_7 = \left|\bar{\mathbf{g}}_{b,r}\Theta_t\bar{\mathbf{g}}_{r,u3u}\right|^2$, $\xi_8 = \left|\bar{\mathbf{g}}_{b,r}^H\Theta_t\bar{\mathbf{g}}_{b,r}\right|^2$ and $\hat{\zeta} = \bar{\mathbf{g}}_{b,r}^H\Theta_t\bar{\mathbf{g}}_{b,r}$.

Proof: The proof is provided in Appendix D. ■

2) Cell-center, Group $\mathcal{G}_2$: The SINR to detect the UL second strongest user signal can be written as,

$$\gamma_{u_{2u}} = \frac{p_{u_{2u}}A_{u_{2u}}}{\Xi p_{u_{1u}}B_{u_{2u}} + p_{u_{3u}}C_{u_{2u}} + P_b D_{u_{2u}} + V + \sigma_b^2} \tag{26}$$

where $A_{u_{2u}} = \left|\sqrt{l_{b,u_{2u}}^{-m}}h_{b,u_{2u}}\right|^2$, $B_{u_{2u}} = \left|\sqrt{l_{b,u_{1u}}^{-m}}h_{b,u_{1u}}\right|^2$ $C_{u_{2u}} = l_{b,r}^{-m}l_{r,u_{3u}}^{-m}\left|\mathbf{g}_{b,r}\Theta_t\mathbf{g}_{r,u_{3u}}\right|^2$, $D_{u_{2u}} = l_{b,r}^{-m}l_{b,r}^{-m}\left|\mathbf{g}_{b,r}^H\Theta_t\mathbf{g}_{b,r}\right|^2$.

Theorem 5. *The ergodic rate of the UL second strongest user in a cluster can be evaluated by*

$$\mathcal{E}\left[R_{u_{2u}}\right] \approx \frac{1}{M_u}\log_2\left(1+\frac{p_{u_{2u}}\chi_{u_{2u}}}{\Xi p_{u_{1u}}\chi_{u_{1u}}+p_{u_{3u}}\omega_{b,r}^{r,u_{3u}}y_{2_{u_{2u}}}+P_b y_{3_{u_{2u}}}+V+\sigma_b^2}\right) \tag{27}$$

where $\chi_{u_{iu}}$, $y_{2_{u_{2u}}} = y_{2_{u1_{2u}}}$ and $y_{3_{u_{2u}}} = y_{3_{u_{1u}}}$ are defined in (23)-(25).

Proof: The proof is provided in Appendix D. ■

3) Cell-edge, Group $\mathcal{G}_3$: The SINR to detect the UL cell-edge user signal can be written as,

$$\gamma_{u_{3u}} = \frac{p_{u_{3u}} A_{u_{3u}}}{\Xi p_{u_{1u}} B_{u_{3u}} + \Xi p_{u_{2u}} C_{u_{3u}} + P_b D_{u_{3u}} + V + \sigma_b^2} \quad (28)$$

where $A_{u_{3u}} = l_{b,r}^{-m} l_{r,u_{3u}}^{-m} \left|\mathbf{g}_{b,r} \Theta_t \mathbf{g}_{r,u_{3u}}\right|^2$,
$B_{u_{3u}} = \left|\sqrt{l_{b,u_{1u}}^{-m}} h_{b,u_{1u}}\right|^2$, $C_{u_{3u}} = \left|\sqrt{l_{b,u_{2u}}^{-m}} h_{b,u_{2u}}\right|^2$
$D_{u_{3u}} = l_{b,r}^{-m} l_{b,r}^{-m} \left|\mathbf{g}_{b,r}^H \Theta \mathbf{g}_{b,r}\right|^2$.

Theorem 6. *The ergodic rate of the UL weakest user in a cluster can be evaluated by*

$$\mathcal{E}\left[R_{u_{3u}}\right] \approx \frac{1}{M_u} \log_2 \left(1 + \frac{p_{u_{3u}} \omega_{b,r}^{r,u_{3u}} x_{1_{u_{3u}}}}{\Xi p_{u_{1u}} \chi_{u_{1u}} + \Xi p_{u_{2u}} \chi_{u_{2u}} + P_b y_{3_{u_{3u}}} + V + \sigma_b^2}\right) \quad (29)$$

where $\chi_{u_{iu}}$, $x_{1_{u_{3u}}} = y_{2_{u_{1u}}}$, and $y_{3_{u_{3u}}} = y_{3_{u_{1u}}}$ are defined in (23)-(25).

Proof: The proof is in Appendix D. ■

C. Design Insights

From the ergodic rate expressions provided in the Theorems we can observe the following. The DL and UL ergodic rates depend essentially on the network geometry, i.e., the radius of the cell center and cell edge areas, the locations of the BS, STAR-RIS and the users, the transmit powers, and number of the STAR-RIS elements. Therefore, for a given network topology, the DL and UL performance is highly related to the STAR-RIS amplitudes and the phase shifts, the transmit powers, and the system impairments, i.e., SIC error and the self interference. From Theorems 1, 2 and 3, the DL performance depends essentially on the BS transmit power, P_b, and the STAR-RIS elements. Thus the DL performance can be enhanced by increasing the BS transmit power and/or number of STAR-RIS units. However, number of STAR-RIS elements has an impact on both the desired signal and the interference, e.g., the nominator and denominator of the SINRs, and thus the STAR-RIS elements should be optimized properly to enhance the DL achievable rates. In addition, the DL ergodic rates degrade with increasing the UL users power and the SIC errors. From Theorems 4, 5 and 6, the UL performance depends on the user transmit power, and the STAR-RIS elements. Similarly as in the DL, the UL users power and number of STAR-RIS elements have an impact on both the desired signal and the interference, and thus these parameters should be optimized to enhance the UL performance. In addition, the UL ergodic rates degrade with increasing the BS transmit power and the self interference at the BS.

In summary, we can conclude that, the optimal STAR-RIS design plays a crucial role in achieving high data rates. In addition, as it is commonly observed in FD communications, higher transmit power generally leads to a stronger signal quality. However, increasing the transmit power can also potentially result in higher interference levels. Thus, the transmit power levels should also be adjusted to optimize the DL and UL performance.

V. Systems Design

As we can observe from the Theorems, the ergodic rates depend on the amplitudes and phase shifts of the STAR-RIS elements and also on the transmit powers. Thus, these parameters are optimized in this Section to maximize the weighted sum rate of a DL cluster and an UL cluster.

A. Simultaneous Amplitudes and Phase-Shifts Optimization

The total weighted sum-rate at a given time slot and power transmission can be optimized by considering the following problem

$$\begin{aligned} \max_{\rho,\theta} \sum_{i=1}^{3} \varpi_{u_{id}} \bar{R}_{u_{id}} + \sum_{i=1}^{3} \varpi_{u_{iu}} \bar{R}_{u_{iu}} \\ \text{s.t } (\rho_n^r) + (\rho_n^t) = 1, \forall n \in N \\ \rho_n^k \geq 0, \left|\theta_n^k\right| = 1, \forall n \in N \end{aligned} \quad (30)$$

where $\varpi_{u_{id}}$ and $\varpi_{u_{iu}}$ are the weighting factors, which signify the priority assigned to each user and $\theta = [\theta^t, \theta^r]$, $\rho = [\rho^t, \rho^r]$. Problem (31) is non-convex and, coupled among the optimization variables which are the amplitudes, ρ, and the phase shifts, θ, for transmission and reflection. The problem is a constrained optimization problem where the objective function is differentiable with respect to the optimization variables (θ,ρ). For the development of an efficient algorithm to solve (31), we notice that the feasible sets Φ (for θ) and Q (for ρ) are simple in the sense that their projection operators can be obtained in closed-form [35] [36]. This motivates us to apply the PGAM to optimize θ and ρ. PGAM results in low computational complexity, being dependent on simple mathematical operations in terms of closed-form expressions. It iteratively updates the optimization variables by taking gradient ascent steps and projecting onto the feasible set. For this method to converge, first the objective function is differentiable with respect to θ and ρ, and its gradients are Lipschitz continuous. In addition, the feasible sets Φ and Q are compact, ensuring that the projection operations remain well-posed [35] [36]. PGAM's convergence is well-established. The projection step ensures that each iteration remains within the feasible set, which is crucial for constrained optimization problems. Thus the convergence of the proposed algorithm to a maximum can be guaranteed because it is bounded by the constraint. To apply the PGAM, first we define $\sum_{i=1}^{3} \varpi_{u_{id}} \bar{R}_{u_{id}} + \varpi_{u_{iu}} \sum_{i=1}^{3} \bar{R}_{u_{iu}} = f(\theta, \rho)$, $\Phi = \left\{\theta^t \in \mathbb{C}^{N\times 1}, \theta^r \in \mathbb{C}^{N\times 1} \,||\theta^r| = |\theta^t| = 1\right\}$, and
$Q = \left\{\rho^t \in \mathbb{C}^{N\times 1}, \rho^r \in \mathbb{C}^{N\times 1},\right.$
$\left.(\rho_i^t) + (\rho_i^r) = 1, \rho_i^t \geq 0, \rho_i^r \geq 0\right\}$. Then, we calculate the gradients of $f(\theta, \rho)$ with respect to θ, $\nabla_\theta f\left(\theta^i, \rho^i\right)$, and ρ^i, $\nabla_\rho f\left(\theta^i, \rho^i\right)$. Next we update the phases and amplitudes at each iteration using the expressions, $\theta^{i+1} = \left(\theta^i + \nu_i \nabla_\theta f\left(\theta^i, \rho^i\right)\right)$ and $\rho^{i+1} = \left(\rho^i + \vartheta_i \nabla_\rho f\left(\theta^i, \rho^i\right)\right)$, respectively, where ν_i and ϑ_i are the step sizes. Then we project them onto Φ and Q. The overall steps of the algorithm is summarized in Algorithm 2.

Algorithm 2 Optimization Algorithm.

Input: Set the maximum number of iterations $\mathscr{W}$ and a tolerance $\mu > 0$.
Initialize $\rho^{(0)}$, and $\theta^{(0)}$, and set the step sizes.
for $i = 0$ to $\mathscr{W}$ do
Evaluate: $f\left(\theta^{(i)}, \rho^{(i)}\right)$, then $\nabla_\theta f\left(\theta^{(i)}, \rho^{(i)}\right)$, and $\nabla_\rho f\left(\theta^{(i)}, \rho^{(i)}\right)$
Update: $\theta^{i+1} = \theta^i + \nu \nabla_\theta f\left(\theta^i, \rho^i\right)$ and $\rho^{i+1} = \rho^i + \vartheta \nabla_\rho f\left(\theta^i, \rho^i\right)$
Evaluate: $f\left(\theta^{(i+1)}, \rho^{(i+1)}\right)$
Until Convergence: $f\left(\theta^{(i+1)}, \rho^{(i+1)}\right) - f\left(\theta^{(i)}, \rho^{(i)}\right) < \mu$.

Complexity Analysis of Algorithms 2: The analytical expressions for the ergodic rates paved the way to apply the PGAM algorithm [35] [36]. This algorithm has low complexity, because it depends only on simple operations in terms of closed-form expressions. Specifically, the Algorithm requires the first-order information, i.e., the objective and its gradient values. Each term requires $\mathcal{O}\left(N^2\right)$ complex multiplications, thus the ergodic rate requires $\mathcal{O}\left(KN^2\right)$iterations. The same number of iterations is required by $\nabla_\theta f\left(\theta, \rho\right)$, and $\nabla_\rho f\left(\theta, \rho\right)$. Obviously, it depends on fundamental system parameters, i.e., K, and N but with a higher dependence on N.

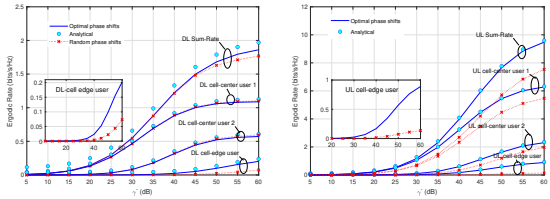
Convergence Analysis of Algorithm 2: The convergence is guaranteed because it iteratively maximizes the rates subject to specific constraints. The gradients $\nabla_\theta f\left(\theta, \rho\right)$, and $\nabla_\rho f\left(\theta, \rho\right)$ are Lipschitz continuous over the feasible set as they comprise basic functions. Let L_θ and L_ρ be the Lipschitz constant of $\nabla_\theta f\left(\theta, \rho\right)$ and $\nabla_\rho f\left(\theta, \rho\right)$, respectively. It holds that

$$f\left(\mathbf{x}, \mathbf{y}\right) \geq f\left(\theta, \rho\right) + \left\langle \nabla_\theta f\left(\theta^i, \rho^i\right), \mathbf{x} - \theta \right\rangle - \frac{1}{L_{\max}} \left\| \mathbf{x} - \theta \right\|_2^2$$

$$+ \left\langle \nabla_\rho f\left(\theta^i, \rho^i\right), \mathbf{y} - \rho \right\rangle - \frac{1}{L_{\max}} \left\| \mathbf{y} - \rho \right\|_2^2 \qquad (31)$$

where $L_{\max} = \max\left(L_\theta, L_\rho\right)$, $\langle \mathbf{a}, \mathbf{b} \rangle = 2\mathrm{Re}\left(\mathbf{a}^H \mathbf{b}\right)$ for complex-values of $\mathbf{a}$ and $\mathbf{b}$ and $\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}^H \mathbf{b}$, for non-complex values of $\mathbf{a}$ and $\mathbf{b}$. Thus, the line search procedure of Algorithm 1 terminates in finite iterations since the condition in Algorithm 1 must be satisfied. Also, due to the line search we automatically have an increasing sequence of objectives, i.e., $f\left(\theta^{(i+1)}, \rho^{(i+1)}\right) \geq f\left(\theta^{(i)}, \rho^{(i)}\right)$. Since the feasible sets Φ and Q are compact, $f\left(\theta^{(i)}, \rho^{(i)}\right)$ must converge.

1) Sub-optimal Design: The sub-optimal solution is based on the proposed NOMA clustering scheme, where in each slot the BS transmits signals to a DL cluster and receives signals from an UL cluster. In such systems the cell-edge users have the highest priority in the STAR-RIS design. Therefore, the phase shifts can be aligned to the cell-edge users' channels. The phase shifts can be presented as, $\phi_n^k = -2\pi \frac{\iota}{\lambda}\left(c_n \upsilon_{b,u_{2k}} + f_n l_{b,u_{2k}}\right), k \in \{t, r\}$ where $\upsilon_{u_{2k}} = \sin\varphi_{b,r}^a \sin\varphi_{b,r}^e - \sin\varphi_{r,u_{2k}}^a \sin\varphi_{r,u_{2k}}^e$, $l_{u_{2k}} = \cos\varphi_{b,r}^e - \cos\varphi_{r,u_{2k}}^e$, and $\varphi_{i,j}^a, \varphi_{i,j}^e$ are the azimuth and elevation angles of arrival from node i to node j, λ is the wavelength, ι is the elements spacing, and $c_n = (n-1) \bmod \sqrt{N}$, $f_n = \frac{n-1}{\sqrt{N}}$. Also, the amplitudes ρ_n^r and ρ_n^t can be calculated using the ergodic rate expressions based on the target data rates.



(a) DL rates versus transmit SNR,$\bar{\gamma}$, for random and optimal phase shifts.

(b) UL rates versus transmit SNR,$\bar{\gamma}$, for random and optimal phase shifts.

Figure 3: Ergodic rates versus transmit SNR,$\bar{\gamma}$, for different phase shifts.

B. Power Allocation

The optimal power allocation scheme can be obtained by solving the problem

$$\max_{\mathbf{p}} \sum_{i=1}^{2} f\left(\mathbf{p}\right)$$

$$(C.1) \quad \text{s.t} \sum_{i=1}^{3} P_{b_i} \leq P_b, P_{b_i} > 0, 0 < p_{u_{iu}} \leq p_{u_m}$$

$$(C.2) \quad \bar{R}_{u_{id}} \geq \hat{R}_{d_i}, \bar{R}_{u_{iu}} \geq \hat{R}_{u_i}, i \in \{1, 2, 3\}$$

$$(C.3) \quad R_{u_{kd} \to u_{id}} \geqslant R_{u_{id}}, k > i \qquad (32)$$

where $f\left(\mathbf{p}\right) = \sum_{i=1}^{3} \varpi_{u_{id}} \bar{R}_{u_{id}} + \sum_{i=1}^{3} \varpi_{u_{iu}} \bar{R}_{u_{iu}}$, $\mathbf{p} = \left[p_{u_1}, p_{u_2}, p_{u_3}, P_{b_1}, P_{b_2}, P_{b_3}\right]^T$, $P_b = \sum_{i=1}^{3} P_{b_i}$ and p_{u_m} is the maximum transmission power at a user. The optimal solution of the problem in (32) can be obtained by using some complicated techniques such as monotonic optimization, and block coordinate descent (BCD) iterative algorithms which are described in [37], [38], the details are omitted here due to the paper length limitation.

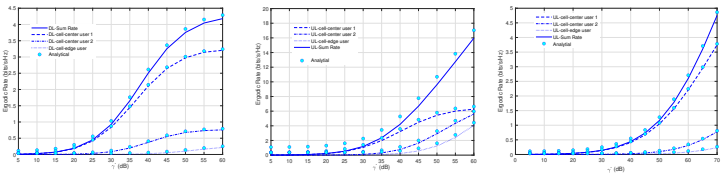
As a simple sub-optimal solution, the minimum power required for each user can be obtained by achieving the minimum data rate requirement of each user, i.e., satisfying $(C.2)$ in (32) with equality, $\bar{R}_{u_{id}} = \hat{R}_{d_i}$, $\bar{R}_{u_{iu}} = \hat{R}_{u_i}$. Thus, using the expressions in Theorems 1-6, the minimum required power values can be obtained by solving all the equality equations together.

VI. Numerical Results

In this section simulation and numerical results are presented to demonstrate the effectiveness of the proposed schemes and confirm the accuracy of the analytical expressions. The radius of the cell-center and cell-edge areas are assumed to be $R = 50m$ and $R_r = 30m$ respectively. For simplicity, it is assumed that the users have the same noise variance, σ^2, and thus the DL transmit SNR is defined as $\bar{\gamma} = \frac{P_b}{\sigma^2}$, and the maximum UL user power is $p_{u_m} = 0.1P_b$ [39]. The Rician factors are 3, and the path-loss exponent is $m = 2.7$. In addition, number of the STAR-RIS elements is $N = 10$, number of the DL cell-center users is K_{cd}=6, and the DL cell-edge users is $K_{ed} = 3$, number of the UL cell-center users is K_{cu}=6, and the UL cell-edge users is $K_{eu} = 3$, the self interference parameters are $\beta = 0.001, \lambda = 0.1$ and the SIC error factor is $\Xi = 0.1$.

In Fig. 3 we illustrate the ergodic rates of a DL cluster and an UL cluster versus the transmit SNR, $\bar{\gamma}$. Fig. 3a shows the achievable DL ergodic rates using the optimal and random phase

This article has been accepted for publication in IEEE Transactions on Wireless 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/TWC.2025.3568772

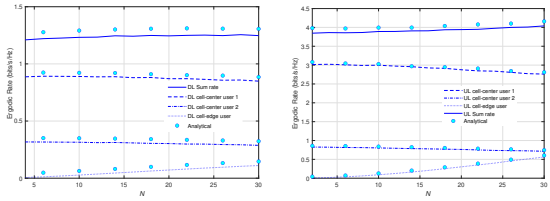


(a) DL rates versus transmit SNR$\bar{\gamma}$, with ideal SIC $\Xi = 0$.

(b) UL rates versus transmit SNR$\bar{\gamma}$, with ideal SIC $\Xi = 0$.

(c) UL rates versus the transmit SNR, $\bar{\gamma}$, when $\beta = 1, \lambda = 0.4$.

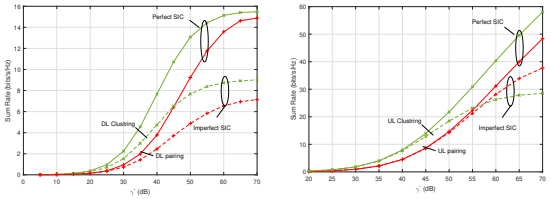
Figure 4: Ergodic rates versus transmit SNR,$\bar{\gamma}$, with different values of the SIC error factor and self interference.



(a) DL rates versus number of the STAR-RIS elements N.

(b) UL rates versus number of the STAR-RIS elements N.

Figure 6: Ergodic rates versus number of the STAR-RIS elements N.



(a) DL sum-rates of user-clustering and user-paring schemes versus the transmit SNR, $\bar{\gamma}$.

(b) UL sum-rates of user-clustering and user-paring schemes versus the transmit SNR, $\bar{\gamma}$.

Figure 5: Sum-rates of user-clustering and user-paring schemes versus the transmit SNR, $\bar{\gamma}$, for perfect and imperfect SIC.

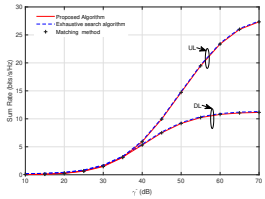


Figure 7: Sum-rate versus transmit SNR for different optimization schemes.

shifts, while the UL ergodic rates with the optimal and random phase shifts are presented in Fig. 3b. Firstly, it is clear from these results that the DL and UL rates enhance with increasing the transmit SNR. In addition, the cell-edge users achieve higher data rates in the optimal phase shifts case than that in the random phase shifts case. Interestingly enough, the performance of the UL cell-center users improves in the random phase shifts case, due to reducing the interference power caused by the UL cell-edge users.

To show the impact of the system impairments on the users' performance, in Fig. 4 we plot the DL and UL ergodic rates versus the transmit SNR, $\bar{\gamma}$, for ideal SIC and high self interference at the FD-BS. Fig. 4a and Fig. 4b depict the DL and UL achievable rates with ideal SIC $\Xi = 0$, respectively, while Fig. 4c presents the UL ergodic rates when the variance of the residual self interference at the BS is high with, $\beta = 1, \lambda = 0.4$. It can be observed from Figs. 4a and 3a that the imperfect SIC results in deteriorating the achievable rates of the DL cell-center users as their performance relies on the SIC detection scheme. In addition, from Fig. 4b and Fig. 3b, we can see that, the performance of the UL cell-center user 2 and the UL cell-edge user enhance greatly with reducing the SIC error factor. Also, comparing Fig. 4c and Fig. 3b, we can observe that, high residual self interference at the BS causes deep degradation of the UL users' performance.

To compare NOMA user clustering scheme with NOMA pairing scheme considered in [20], we plot in Fig. 5 the DL and UL sum rates of the two schemes versus the transmit SNR, $\bar{\gamma}$, for perfect and imperfect SIC, $\Xi = 0$ and $\Xi = 0.1$,e.g., 0% and 10% of the power remains as interference. The power in these results is allocated between the users to achieve the target data rates of the cell-edge users, and the remaining power is divided between the cell-center users to enhance the total sum rate. It is worth mentioning that, in NOMA user clustering scheme there are 3 different DL and UL clusters, while in NOMA pairing scheme, there are 5 different DL and UL clusters. From Fig. 5a, in DL mode NOMA user clustering scheme achieves higher sum rates than NOMA pairing scheme. On the other hand, from Fig. 5b in UL mode NOMA user clustering scheme outperforms NOMA paring scheme in the perfect SIC case, and at low SNR values in the imperfect SIC case. However, in a high SNR regime NOMA pairing scheme can achieve higher sum-rates. The sum-rate of a cluster is controlled by the performance of the cell-center users. In DL increasing SIC error hurts the performance of the cell-center users in the two schemes, while in UL mode increasing SIC error degrades the performance of the cell-edge users which can be controlled by the power allocation scheme.

Fig. 6, illustrates the ergodic rates versus number of the STAR-RIS elements N when $\bar{\gamma} = 40$ dB. Fig. 6a shows that the performance of the DL cell-edge user is enhanced with increasing number of the STAR-RIS elements, N, while the cell-center users performance degrades with increasing N. This is because increasing number of the elements leads to increase the interference power caused by the UL cell-edge user. On the other hand, Fig. 5b demonstrates that the performance of the UL cell-edge user is enhanced greatly with increasing N, which is not the case for the UL cell-center users where adding more elements results in increasing the interference power caused by the UL cell-edge user.

Fig. 7 shows the total sum-rate for the proposed scheme, matching theory and the exhaustive search algorithm to find the optimal clustering for 6 DL and 6 UL cell-center users and 3 DL and 3 UL cell edge users, e.g., K_{cd}=6, K_{ed}=3, K_{cu}=6, and K_{eu}=3. As we can notice from the figure the gap is tiny for the considered scenario. The exhaustive search algorithm is computationally intensive compared to our proposed scheme which is much simpler.

VII. Conclusions

This work considered a STAR-RIS-assisted FD NOMA communication system, where the STAR-RIS is implemented at the cell-edge region to assist the cell-edge users. Firstly, new user clustering schemes for DL and UL transmissions were presented. Then, closed-form analytical expressions of the DL and UL ergodic sum rates have been derived. In addition, the optimal amplitudes and phase-shifts of the STAR-RIS elements that maximize the total sum-rate were obtained. Moreover, a power allocation scheme of the DL and UL communications was studied. The results in this work showed that increasing the transmitted SNR always improves the achievable rates, and the performance of the cell-edge users can be enhanced by using a large number of STAR-RIS elements. Furthermore, imperfect SIC degrades the achievable rates of the DL cell-center users and the UL cell-edge user, while imperfect SIS reduces the performance of the UL users significantly. In addition, NOMA user clustering scheme can achieve a higher sum rate than NOMA pairing scheme in DL mode, while in UL mode NOMA user clustering scheme outperforms NOMA paring scheme in the perfect SIC case, and at low SNR values in the imperfect SIC case.

Appendix A

By using Jensen inequality, the ergodic rate can be approximated by

$$\mathcal{E}\left[R_{u_{1d}}\right] \approx \log_2\left(1+ \frac{P_{b_1}\mathcal{E}\left\{A_{u_{1d}}\right\}}{\mathcal{E}\left\{\Xi P_b\left(\alpha_2+\alpha_3\right)A_{u_{1d}}+B_{u_{1d}}+C_{u_{1d}}+D_{u_{1d}}\right\}+\sigma^2_{u_{1d}}}\right) \tag{33}$$

Due to the paper length limitation, we refer the reader to [40, Lemma 1] for more details about the error analysis of using Jensen inequality. The first term, $\mathcal{E}\left\{A_{u_{1d}}\right\} = \mathcal{E}\left\{l^{-m}_{b,u_{1d}}\right\} = x_{1_{u_{1d}}}$ can be calculated as follows. The PDF of the users at radius r relative to the BS is $f_d(r) = \frac{2r}{R^2}, \quad 0 \le r \le R$. Thus, the density of the kth user order can be expressed as

$$f_{d_k}(r) = \frac{K_{cd}!}{(k-1)!\,(K_{cd}-1)!}\frac{2(r)}{(R)^2}\left(\frac{r^2}{R^2}\right)^{k-1}\left(1-\frac{r^2}{R^2}\right)^{K_{cd}-k} \tag{34}$$

Then, we can find the average by

$$x_{1_{u_{1d}}} = \frac{2K_{cd}!}{(k-1)!\,(K_{cd}-1)!}\int_0^R \left(1+r_{u_{1d}}\right)^{-m} \times\frac{r}{(R)^2}\left(\frac{r^2}{(R)^2}\right)^{k-1}\left(1-\frac{r^2}{R^2}\right)^{K_{cd}-k} dr_{u_{1d}} \tag{35}$$

The solution of (35) is presented in (10). The second term, $\mathcal{E}\left\{B_{u_{1d}}\right\} = \mathcal{E}\left\{l^{-m}_{u_{1d},u_{1u}}\right\} = y_{1_{u_{1d}}}$ can be calculated as follows. Considering the distribution of the distance between two random points inside a circle in [41], we can write

$$y_{1_{u_{1d}}} = \int_0^{2R}\left(1+r_{u_{1d},u_{1u}}\right)^{-m}\frac{4r_{u_{1d},u_{1u}}}{\pi R^2} \times\left(\cos^{-1}\left(\frac{r_{u_{1d},u_{1u}}}{2R}\right)-\frac{r_{u_{1d},u_{1u}}}{2R}\left(\sqrt{1-\frac{r^2_{u_{1d},u_{1u}}}{4R^2}}\right)\right)dr_{u_{1d},u_{1u}} \tag{36}$$

The solution of (36) is presented in (11). The third term, $\mathcal{E}\left\{C_{u_{1d}}\right\}$ can be calculated by following similar steps as in (36). The last term $\mathcal{E}\left\{D_{u_{1d}}\right\} = p_{u_{3u}}\mathcal{E}\left\{l^{-m}_{r,u3u}\right\}\mathcal{E}\left\{l^{-m}_{r,u1d}\right\}\mathcal{E}\left\{\left|g_{r,u_{1d}}\Theta_t g_{r,u_{3u}}\right|^2\right\}$ can be evaluated as follows. The PDF of the weak user location at radius r_r relative to the RIS is $f_d(r_r) = \frac{2r_r}{R_r^2}, \quad 0 \le r_r \le R_r$. Thus, we can calculate the first average as,

$$\mathcal{E}\left\{l^{-m}_{r,u3u}\right\} = \frac{2K_{eu}!}{(k-1)!\,(K_{eu}-k)!}\int_0^{R_r}\left(1+r_{r,u3u}\right)^{-m} \times\frac{r}{(R_r)^2}\left(\frac{r^2}{(R_r)^2}\right)^{k-1}\left(1-\frac{r^2}{R_r^2}\right)^{K_{eu}-k} dr_{r,u3u} \tag{37}$$

which can be found as in (38). By invoking the distribution of the distance between a random point inside a circle and a fixed point outside the circle provided in [41], we can write

$$\mathcal{E}\left\{l^{-m}_{r,u_{1d}}\right\} = \int_{r_1}^{r_1+2R} r^{-m}_{r,u1d}\frac{2r_{r,u_{1d}}}{\pi R^2}\cos^{-1}\left(\frac{1}{r_{r,u_{1d}}}\left(r_1+\frac{\left(r^2_{r,u_{1d}}-r_1^2\right)}{2\left(R+r_1\right)}\right)\right)dr_{r,u_{1d}} \tag{39}$$

where $r_1 = d_{b,r} - R$ is the distance from RIS to the circle boundary, $r_1 \le r_{r,u_{1d}} \le r_1 + 2R$. Applying Gaussian Quadrature rules we can get

$$\mathcal{E}\left\{l^{-m}_{r,u1d}\right\} = \sum_{j=1}^{C}\mathrm{H}_j\left(1+\left(R\,r_j+R\right)\right)^{-m}\frac{2\left(R\,r_j+R\right)}{\pi R^2} \times\cos^{-1}\left(\frac{1}{\left(R\,r_j+R\right)}\left(r_1+\frac{\left(\left(R\,r_j+R\right)^2-r_1^2\right)}{2\left(R+r_1\right)}\right)\right) \tag{40}$$

After removing the zero expectation terms, the last expectation can be written as in (41). Now, the first term in (41) is

$$\mathscr{E}\left|\bar{\mathbf{g}}_{r,u1d}\Theta\bar{\mathbf{g}}_{r,u_{3u}}\right|^2 = = \left|\sum_{n=1}^{N} a_{N,n}\left(\psi^a_{r,u_{3u}},\psi^e_{r,u_{3u}}\right)\rho^k_n e^{j\phi^k_n}a_{N,n}\left(\psi^a_{r,u_{1d}},\psi^e_{r,u_{1d}}\right)\right|^2 = \xi_1 \tag{42}$$

Similarly, the second term,

$$\mathscr{E}\left|\bar{\mathbf{g}}_{r,u1d}\Theta\tilde{\mathbf{g}}_{r,u_{3u}}\right|^2 = \sum_{n=1}^{N}\left|\rho^k_n\right|^2 + \mathscr{E}\left\{\sum_{n_1=1}\sum_{n_2\neq n_1}^{N}\left(a_{Nn_1}\left(\psi^a_{r,u_{1d}},\psi^e_{r,u_{1d}}\right)\rho^k_{n_1}e^{j\phi^k_{n_1}}\left[\tilde{\mathbf{g}}_{r,u_{3u}}\right]_{n_1}\right)\left(a_{Nn_2}\left(\psi^a_{r,u_{1d}},\psi^e_{r,u_{1d}}\right)\rho^k_{n_2}e^{j\phi^k_{n_2}}\left[\tilde{\mathbf{g}}_{r,u_{3u}}\right]_{n_2}\right)^H\right\} = \sum_{n=1}^{N}\left|\rho^k_n\right|^2 \tag{43}$$

$$y_{2_{u_{1d}}} = \left(2a_{1_{u3u}}\Gamma\left(k_{eu_3}\right)\mathcal{H}\left(\left\{k_{eu_3}, \frac{1+m}{2}, \frac{m}{2}\right\}, \left\{\frac{1}{2}, 1+K_{eu}\right\}, R_r^2\right)\right.$$
$$\left. -mR_r a_{1_{u3u}}\Gamma\left(\frac{1}{2}+k_{eu_3}\right)\mathcal{H}\left(\left\{\frac{1}{2}+k_{eu_3}, \frac{1+m}{2}, \frac{2+m}{2}\right\}, \left\{\frac{3}{2}, \frac{3}{2}+K_{eu}\right\}, R_r^2\right)\right) \tag{38}$$

$$\mathscr{E}\left\{\left|\mathbf{g}_{r,u1d}\Theta\mathbf{g}_{r,u_{3u}}\right|^2\right\} = \frac{\kappa_{r,u1d}}{\kappa_{r,u1d}+1}\frac{\kappa_{r,u_{3u}}}{\kappa_{r,u_{3u}}+1}\mathscr{E}\left|\bar{\mathbf{g}}_{r,u1d}\Theta\bar{\mathbf{g}}_{r,u_{3u}}\right|^2 + \frac{\kappa_{r,u1d}}{\kappa_{r,u1d}+1}\frac{1}{\kappa_{r,u_{3u}}+1}\mathscr{E}\left|\bar{\mathbf{g}}_{r,u1d}\Theta\tilde{\mathbf{g}}_{r,u_{3u}}\right|^2$$
$$+\frac{\kappa_{r,u_{3u}}}{\kappa_{r,u_{3u}}+1}\frac{1}{\kappa_{r,u1d}+1}\mathscr{E}\left|\tilde{\mathbf{g}}_{r,u1d}\Theta\bar{\mathbf{g}}_{r,u_{3u}}\right|^2 + \frac{1}{\kappa_{r,u1d}+1}\frac{1}{\kappa_{r,u_{3u}}+1}\mathscr{E}\left|\tilde{\mathbf{g}}_{r,u1d}\Theta\tilde{\mathbf{g}}_{r,u_{3u}}\right|^2 \tag{41}$$

The other terms,

$$\mathscr{E}\left|\tilde{\mathbf{g}}_{r,u1d}\Theta\bar{\mathbf{g}}_{r,u_{3u}}\right|^2 = \sum_{n=1}^{N}\left|\rho_n^k\right|^2 +$$

$$\mathscr{E}\left\{\sum_{n_1=1}^{N}\sum_{n_2\neq n_1}^{N}\left(\left[\tilde{\mathbf{g}}_{r,u1d}\right]_{n_1}\rho_{n_1}^k e^{j\phi_{n_1}^k}a_{Nn1}\left(\psi_{r,u_{3u}}^a, \psi_{r,u_{3u}}^e\right)\right)\right.$$

$$\left.\left(\left[\tilde{\mathbf{g}}_{r,u1d}\right]_{n_2}\rho_{n_1}^k e^{j\phi_{n_1}^k}a_{Nn_2}\left(\psi_{r,u_{3u}}^a, \psi_{r,u_{3u}}^e\right)\right)^H\right\} = \sum_{n=1}^{N}\left|\rho_n^k\right|^2 \tag{44}$$

and

$$\mathscr{E}\left|\tilde{\mathbf{g}}_{r,u1d}\Theta\tilde{\mathbf{g}}_{r,u_{3u}}\right|^2 = \sum_{n=1}^{N}\left|\rho_n^k\right|^2 \tag{45}$$

APPENDIX B

By using Jensen inequality, the average can be derived as follows. The first term, $\mathcal{E}\left\{A_{u_{2d}}\right\} = \mathcal{E}\left\{l_{b,u_{2d}}^{-m}\right\}$ can be calculated as

$$\mathcal{E}\left\{l_{b,u_{2d}}^{-m}\right\} = \frac{2K_{cd}!}{\left(k_{cd_2}-1\right)!\left(K_{cd}-1\right)!}\int_0^R\left(1+r_{u_{2d}}\right)^{-m}$$

$$\times\frac{(r)}{(R)^2}\left(\frac{r^2}{R^2}\right)^{k_{cd_2}-1}\left(1-\frac{r^2}{R^2}\right)^{K_{cd}-k_{cd_2}}dr_{u_{2d}} \tag{46}$$

which can be found as in (15). The derivation of $\mathcal{E}\left\{B_{u_{2d}}\right\} = \mathcal{E}\left\{l_{u_{2d},u_{1u}}^{-m}\right\}, \mathcal{E}\left\{C_{u_{2d}}\right\} = \mathcal{E}\left\{l_{u_{2d},u_{2u}}^{-m}\right\}$ can be obtained using the distribution of the distance between two random points inside a circle as $\mathcal{E}\left\{l_{u_{2d},u_{iu}}^{-m}\right\} = \int_0^{2R}\left(1+r_{u_{2d},u_{iu}}\right)^{-m}\frac{4r_{u_{2d},u_{iu}}}{\pi R^2}\left(\cos^{-1}\left(\frac{r_{u_{2d},u_{iu}}}{2R}\right)\right.$ $\left.-\frac{r_{u_{2d},u_{iu}}}{2R}\left(\sqrt{1-\frac{r_{u_{2d},u_{iu}}^2}{4R^2}}\right)\right)dr_{u_{2d},u_{iu}}$ which can by found as in (11). In addition, $\mathcal{E}\left\{D_{u_{2d}}\right\} = \mathcal{E}\left\{l_{r,u3u}^{-m}l_{r,u2d}^{-m}\left|\mathbf{g}_{r,u_{2d}}\Theta_t\mathbf{g}_{r,u_{3u}}\right|^2\right\}$, where $\mathcal{E}\left\{l_{r,u2d}^{-m}\right\}$ and $\mathcal{E}\left\{\left|\mathbf{g}_{r,u_{2d}}\Theta_t\mathbf{g}_{r,u_{3u}}\right|^2\right\}$ can be derived by following similar steps as in (39), and (41), respectively.

APPENDIX C

To find the ergodic rate using Jensen inequality, we need to derive the average of the following terms. To find $\mathcal{E}\left\{l_{r,uiu}^{-m}\right\}$where $i = 1, 2$, we use the distribution of the distance between a random point inside a circle and a fixed point outside the circle provided in [41] as $\mathcal{E}\left\{l_{r,uiu}^{-m}\right\} = \int_{r_1}^{r_1+2R} r_{r,uiu}^{-m}\frac{2r_{r,uiu}}{\pi R^2}\cos^{-1}\left(\frac{1}{r_{r,uiu}}\left(r_1+\frac{\left(r_{r,uiu}^2-r_1^2\right)}{2(R+r_1)}\right)\right)dr_{r,uiu}$. Applying Gaussian Quadrature rules we can get the average as in (40). The terms $\mathcal{E}\left\{l_{r,u3d}^{-m}\right\}$ and $\mathcal{E}\left\{\left|\mathbf{g}_{r,u3d}\Theta\mathbf{g}_{b,r}\right|^2\right\}$, $\mathcal{E}\left\{\left|\mathbf{g}_{r,u3d}\Theta_k\mathbf{g}_{u_{1u},r}\right|^2\right\}$, $\mathcal{E}\left\{\left|\mathbf{g}_{r,u3d}\Theta\mathbf{g}_{u_{2u},r}\right|^2\right\}$, $\mathcal{E}\left\{\left|\mathbf{g}_{r,u_{3d}}\Theta_r\mathbf{g}_{r,u_{3u}}\right|^2\right\}$can be derived by following the steps as in (37) and (41).

APPENDIX D

To derive the ergodic rate using Jensen inequality, we should derive the average of the following. The terms, $\mathcal{E}\left\{l_{b,u_{1u}}^{-m}\right\}$ and $\mathcal{E}\left\{l_{b,u_{2u}}^{-m}\right\}$ can be calculated as $\mathcal{E}\left\{l_{b,u_{iu}}^{-m}\right\} = \frac{2K_{cu}!}{\left(k_{cu_i}-1\right)!\left(K_{cu}-1\right)!}\int_0^R r_{u_{iu}}^{-m}\frac{r}{(R)^2}\left(\frac{r^2}{(R)^2}\right)^{k-1}\left(1-\frac{r^2}{R^2}\right)^{K_{cu}-k_{cu_i}}dr_{u_{iu}}$ which can be found as in (23) when $i = 1$ and (24) for $i = 2$. Similarly, the terms, $\mathcal{E}\left\{\left|\mathbf{g}_{b,r}\Theta\mathbf{g}_{r,u3u}\right|^2\right\}$ and $\mathcal{E}\left\{\left|\mathbf{g}_{b,r}\Theta_k\mathbf{g}_{b,r}^H\right|^2\right\}$ can be calculated as in the previous derivations in (41).

REFERENCES

[1] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2450–2525, 2020.

[2] C. Pan, H. Ren, K. Wang, J. F. Kolb, M. Elkashlan, M. Chen, M. Di Renzo, Y. Hao, J. Wang, A. L. Swindlehurst, X. You, and L. Hanzo, "Reconfigurable intelligent surfaces for 6G systems: Principles, applications, and research directions," *IEEE Communications Magazine*, vol. 59, no. 6, pp. 14–20, 2021.

[3] A. Salem, K.-K. Wong, and C.-B. Chae, "Impact of phase-shift error on the secrecy performance of uplink ris communication systems," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2023.

[4] Y. Liu, X. Mu, J. Xu, R. Schober, Y. Hao, H. V. Poor, and L. Hanzo, "STAR: Simultaneous transmission and reflection for 360Â° coverage by intelligent surfaces," *IEEE Wireless Communications*, vol. 28, no. 6, pp. 102–109, 2021.

[5] X. Mu, Y. Liu, L. Guo, J. Lin, and R. Schober, "Simultaneously transmitting and reflecting (STAR) RIS aided wireless communications," *IEEE Transactions on Wireless Communications*, vol. 21, no. 5, pp. 3083–3098, 2022.

[6] J. Xu, Y. Liu, X. Mu, R. Schober, and H. V. Poor, "STAR-RISs: A correlated t and r phase-shift model and practical phase-shift configuration strategies," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 5, pp. 1097–1111, 2022.

[7] H. Liu, G. Li, X. Li, Y. Liu, G. Huang, and Z. Ding, "Effective capacity analysis of STAR-RIS-assisted NOMA networks," *IEEE Wireless Communications Letters*, vol. 11, no. 9, pp. 1930–1934, 2022.

[8] Z. Xie, W. Yi, X. Wu, Y. Liu, and A. Nallanathan, "STAR-RIS aided NOMA in multicell networks: A general analytical framework with gamma distributed channel modeling," *IEEE Transactions on Communications*, vol. 70, no. 8, pp. 5629–5644, 2022.

[9] Z. Shi, S. Ma, H. ElSawy, G. Yang, and M. Alouini, "Cooperative harq-assisted noma scheme in large-scale d2d networks," *IEEE Transactions on Communications*, vol. 66, no. 9, pp. 4286–4302, Sep. 2018.

[10] F. Fang, B. Wu, S. Fu, Z. Ding, and X. Wang, "Energy-efficient design of STAR-RIS aided MIMO-NOMA networks," *IEEE Transactions on Communications*, vol. 71, no. 1, pp. 498–511, 2023.

[11] J. Chen and X. Yu, "Ergodic rate analysis and phase design of STAR-RIS aided NOMA with statistical CSI," *IEEE Communications Letters*, vol. 26, no. 12, pp. 2889–2893, 2022.

[12] J. Zuo, Y. Liu, Z. Ding, L. Song, and H. V. Poor, "Joint design for simultaneously transmitting and reflecting (STAR) RIS assisted NOMA systems," *IEEE Transactions on Wireless Communications*, vol. 22, no. 1, pp. 611–626, 2023.

[13] X. Yue, J. Xie, Y. Liu, Z. Han, R. Liu, and Z. Ding, "Simultaneously transmitting and reflecting reconfigurable intelligent surface assisted NOMA networks," *IEEE Transactions on Wireless Communications*, vol. 22, no. 1, pp. 189–204, 2023.

[14] T. Riihonen, S. Werner, and R. Wichman, "Hybrid full-duplex/half-duplex relaying with transmit power adaptation," *IEEE Transactions on Wireless Communications*, vol. 10, no. 9, pp. 3074–3085, 2011.

[15] G. Zheng, I. Krikidis, and B. o. Ottersten, "Full-duplex cooperative cognitive radio with transmit imperfections," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 2498–2511, 2013.

[16] Z. Peng, Z. Zhang, C. Pan, L. Li, and A. L. Swindlehurst, "Multiuser full-duplex two-way communications via intelligent reflecting surface," *IEEE Transactions on Signal Processing*, vol. 69, pp. 837–851, 2021.

[17] P. Guan, Y. Wang, H. Yu, and Y. Zhao, "Joint beamforming optimization for RIS-aided full-duplex communication," *IEEE Wireless Communications Letters*, vol. 11, no. 8, pp. 1629–1633, 2022.

[18] Y. Cai, M.-M. Zhao, K. Xu, and R. Zhang, "Intelligent reflecting surface aided full-duplex communication: Passive beamforming and deployment design," *IEEE Transactions on Wireless Communications*, vol. 21, no. 1, pp. 383–397, 2022.

[19] P. K. Sharma, N. Sharma, S. Dhok, and A. Singh, "RIS-assisted FD short packet communication with non-linear EH," *IEEE Communications Letters*, vol. 27, no. 2, pp. 522–526, 2023.

[20] A. Salem, K.-K. Wong, C.-B. Chae, and Y. Zhang, "STAR-RIS assisted full-duplex communication networks," *arXiv preprint arXiv:2309.15037*, 2023.

[21] K. Wang, W. Liang, Y. Yuan, Y. Liu, Z. Ma, and Z. Ding, "User clustering and power allocation for hybrid non-orthogonal multiple access systems," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 12, pp. 12 052–12 065, 2019.

[22] D. Kudathanthirige and G. A. A. Baduge, "NOMA-aided multicell downlink massive MIMO," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 3, pp. 612–627, 2019.

[23] M. S. Ali, H. Tabassum, and E. Hossain, "Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (NOMA) systems," *IEEE Access*, vol. 4, pp. 6325–6343, 2016.

[24] B. Di, L. Song, and Y. Li, "Sub-channel assignment, power allocation, and user scheduling for non-orthogonal multiple access networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7686–7698, 2016.

[25] M. Liu, J. Zhang, K. Xiong, M. Zhang, P. Fan, and K. B. Letaief, "Effective user clustering and power control for multiantenna uplink NOMA transmission," *IEEE Transactions on Wireless Communications*, vol. 21, no. 11, pp. 8995–9009, 2022.

[26] E. Everett, A. Sahai, and A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes," *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 680–694, 2014.

[27] E. Ahmed and A. M. Eltawil, "All-digital self-interference cancellation technique for full-duplex systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 7, pp. 3519–3532, 2015.

[28] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, I. Chih-Lin, and H. V. Poor, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 185–191, 2017.

[29] M. Zeng, A. Yadav, O. A. Dobre, G. I. Tsiropoulos, and H. V. Poor, "Capacity comparison between MIMO-NOMA and MIMO-OMA with multiple users in a cluster," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2413–2424, 2017.

[30] Y. Liu, Z. Ding, M. Elkashlan, and H. V. Poor, "Cooperative non-orthogonal multiple access with simultaneous wireless information and power transfer," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 938–953, April 2016.

[31] Z. Ding, R. Schober, and H. V. Poor, "A general mimo framework for noma downlink and uplink transmission based on signal alignment," *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 4438–4454, June 2016.

[32] H. Chingoska, Z. Hadzi-Velkov, I. Nikoloska, and N. Zlatanov, "Resource allocation in wireless powered communication networks with non-orthogonal multiple access," *IEEE Wireless Communications Letters*, vol. 5, no. 6, pp. 684–687, Dec 2016.

[33] H. V. Nguyen, V.-D. Nguyen, O. A. Dobre, D. N. Nguyen, E. Dutkiewicz, and O.-S. Shin, "Joint power control and user association for NOMA-based full-duplex systems," *IEEE Transactions on Communications*, vol. 67, no. 11, pp. 8037–8055, 2019.

[34] W. Liang, Z. Ding, Y. Li, and L. Song, "User pairing for downlink non-orthogonal multiple access networks using matching algorithm," *IEEE Transactions on Communications*, vol. 65, no. 12, pp. 5319–5332, 2017.

[35] C. Pan, G. Zhou, K. Zhi, S. Hong, T. Wu, Y. Pan, H. Ren, M. D. Renzo, A. Lee Swindlehurst, R. Zhang, and A. Y. Zhang, "An overview of signal processing techniques for ris/irs-aided wireless systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 5, pp. 883–917, 2022.

[36] A. Papazafeiropoulos, C. Pan, P. Kourtessis, S. Chatzinotas, and J. M. Senior, "Intelligent reflecting surface-assisted mu-miso systems with imperfect hardware: Channel estimation and beamforming design," *IEEE Transactions on Wireless Communications*, vol. 21, no. 3, pp. 2077–2092, 2022.

[37] Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, "Optimal joint power and subcarrier allocation for full-duplex multicarrier non-orthogonal multiple access systems," *IEEE Transactions on Communications*, vol. 65, no. 3, pp. 1077–1091, 2017.

[38] A. Abrardo, M. Moretti, and F. Saggese, "Power and subcarrier allocation in 5G NOMA-FD systems," *IEEE Transactions on Wireless Communications*, vol. 19, no. 12, pp. 8246–8260, 2020.

[39] A. Papazafeiropoulos, P. Kourtessis, and I. Krikidis, "STAR-RIS assisted full-duplex systems: Impact of correlation and maximization," *IEEE Communications Letters*, vol. 26, no. 12, pp. 3004–3008, 2022.

[40] Q. Zhang, S. Jin, K.-K. Wong, H. Zhu, and M. Matthaiou, "Power scaling of uplink massive mimo systems with arbitrary-rank channel means," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 966–981, 2014.

[41] A. M. Mathai, *An Introduction to Geometrical Probability*. Gordon and Breach Science Publishers, 1999.