Rate-Splitting Multiple Access for a Multi-RIS-Assisted Cell-Free Network with Low-Resolution DACs

Mario R. Camana, Zaid Abdullah, Carla E. Garcia, Eva Lagunas and Symeon Chatzinotas
Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, 4365 Luxembourg City, Luxembourg
Email: {mario.camana, zaid.abdullah, carla.garcia, eva.lagunas, symeon.chatzinotas}@uni.lu

***Abstract*—In this paper, we investigate the performance of the rate-splitting multiple access (RSMA) framework in a mmWave cell-free massive multiple-input multiple-output (CF-mMIMO) system assisted by multiple reconfigurable intelligent surfaces (RISs). We consider the practical scenario of low-resolution digital-to-analog converters (DACs) at the distributed access points (APs) to reduce hardware complexity and power consumption. Our main objective is to maximize the minimum rate among the users by jointly optimizing the precoding vectors at each AP, the common rates, and the reflection coefficients of RISs. The resultant non-convex optimization problem is then solved using alternating optimization and successive convex approximation-based methods. Numerical results demonstrate the superior performance of the proposed RSMA-based scheme over traditional methods across several deployment scenarios, with performance gains from RIS deployment notably improved in hotspot scenarios.**

***Index Terms*—Cell-free massive MIMO, rate-splitting multiple access (RSMA), reconfigurable intelligent surface (RIS), digital-to-analog converters (DACs), millimeter wave MIMO.**

I. Introduction

With the expected rapid increase in traffic demand for next-generation wireless communication systems, the roles of massive multiple-input multiple-output (mMIMO) and millimeter-wave (mmWave) technologies, which are already standardized as key wireless transmission technologies in 5G networks, are projected to become increasingly critical. However, intercell interference is one of the main obstacles to improve the capacity in centralized MIMO systems, especially for users located at the cell boundaries. To address this issue, cell-free (CF)-mMIMO systems have been proposed to efficiently manage intercell interference through the coordination of several distributed access points (APs) in the area of interest, thereby enhancing system capacity [1]. In CF-mMIMO systems, the APs simultaneously serve all users within the same time-frequency resource block, with coordination among the APs facilitated by a central processing unit (CPU) via backhaul links. The superior performance of CF-mMIMO over co-located mMIMO has been proven in the literature in terms of energy efficiency [2] and spectral efficiency (SE) [3]. Furthermore, reducing the bit resolution of digital-to-analog converters (DACs) is a well-known solution to reduce hardware complexity and power consumption in mmWave mMIMO systems [4]. The impact of low-resolution DACs in a CF-mMIMO system was investigated in [5].

Moreover, the reconfigurable intelligent surface (RIS) has emerged as a promising technology, particularly for mmWave systems to combat the high channel attenuation and provide additional paths to improve system capacity and coverage [6]. The interplay of RIS in a CF-mMIMO system with multi-antenna APs was investigated in [7] to maximize the sum rate and in [8] to maximize the energy efficiency, demonstrating the performance gains achieved by deploying RISs in such network setups.

Additionally, and when it comes to the wireless access scheme, the rate splitting multiple access (RSMA) has shown to be a promising solution to enhance spectral and energy efficiency in various scenarios involving multi-antenna transmitters [9]-[12]. RSMA is based on splitting the user's message at the transmitter and utilizing successive interference cancellation (SIC) at the receiver, which allows for partial decoding of interference while partially treating it as noise. Initial studies of RSMA in CF-mMIMO systems were conducted in [11] and [12] without considering the deployment of RISs or the impact of low-resolution DACs.

Motivated by the above, we investigate an RSMA framework in a multi-RIS-assisted CF-mMIMO system. The proposed system model consists of several APs equipped with multiple antennas, multiple RISs, and multiple single-antenna users, considering a mmWave system with low-resolution DACs at the APs. To the best of our knowledge, the literature on CF-mMIMO assisted by RISs has not yet investigated the performance improvement and challenges associated with using the RSMA scheme along with the effects of low-resolution DACs. The contributions of the paper are summarized as follows:

- We investigate the maximization of the minimum rate among users in a multi-RIS-assisted CF-mMIMO system with RSMA and low-resolution DACs, subject to the maximum transmission power at each AP.
- Due to the coupling of precoding vectors at the APs and reflection coefficients, we propose an scheme based

This work has been supported by the Smart Networks and Services Joint Undertaking (SNS JU) project TERRAMETA under the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101097101, including top-up funding by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee.

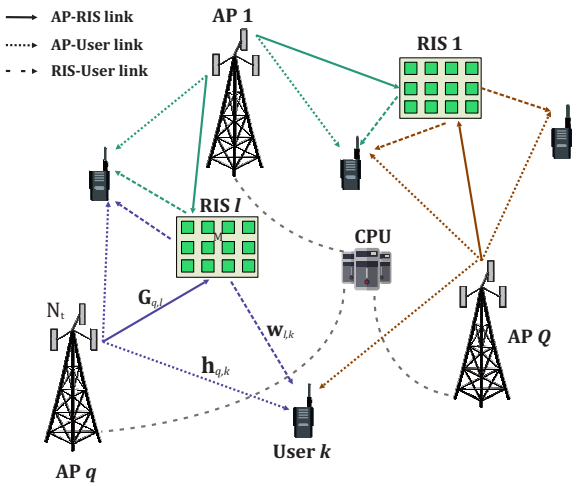

Fig. 1. RSMA CF-mMIMO system assisted by multiple RISs.

on the alternating optimization (AO) algorithm to solve the proposed non-convex problem. The subproblems are addressed using successive-convex approximation (SCA)-based methods by optimizing the precoding vectors for common and private streams, the reflection coefficients of the RISs, and the common rates.

- As comparative methods, we consider the traditional scheme without RSMA, the centralized AP method, the random phase scheme, and the case without deploying RISs. Performance comparison is performed through numerical simulations to demonstrate the superior performance of the proposed scheme based on RSMA over conventional methods. Various deployment scenarios are considered, including variations in the number of transmitters, resolution bits, RISs, and reflecting elements.

Notations: We use lowercase, lowercase bold, and uppercase bold for scalars, vectors, and matrices, respectively. $\mathbb{R}$ and $\mathbb{C}$ denote real and complex values. $\Re()$ denotes the real part of a complex number. $\mathcal{CN}\left(\mu,\sigma^2\right)$ represents a CSCG distribution with mean μ and variance σ^2. Superscripts $\{\cdot\}^T$, and $\{\cdot\}^H$ denote the transpose, and Hermitian transpose, respectively. $\otimes$ represents the Kronecker product, and $\mathbf{I}_N$ is the $N \times N$ identity matrix.

II. System Model and Problem Formulation

We investigate a multi-RIS-assisted CF-mMIMO system with RSMA. The proposed system model considers Q APs with N_t antennas each, K single-antenna users in the same time-frequency resource, and L RISs with M elements each, as illustrated in Fig. 1. We consider one-layer RSMA, dividing the message of the kth user, m_k, into a common message m_k^c and a private message m_k^p. The super-common message is created by grouping all the common messages from the users and encoding them into a common stream s_c using a shared codebook. Each kth private message is encoded independently into the private stream s_k. The precoding vector for the common stream at the qth AP is denoted as $\mathbf{f}_{q,c} \in \mathbb{C}^{N_t \times 1}$ and the precoding vector for the kth private stream at the qth AP is represented by $\mathbf{f}_{q,k} \in \mathbb{C}^{N_t \times 1}$. Then, the transmit signal vector at the qth AP before quantization is given by

$$\mathbf{x}_q = \mathbf{f}_{q,c} s_c + \sum_{k=1}^{K} \mathbf{f}_{q,k} s_k. \tag{1}$$

The low-resolution DAC architecture is deployed at each AP, where we assume that all the DACs in each AP have the same resolution. The distortion due to quantization at the DACs is modeled based on the linear model approximation from [4]. Then, the quantized version of $\mathbf{x}_q$ is represented by

$$\tilde{\mathbf{x}}_q = \lambda_q \mathbf{f}_{q,c} s_c + \lambda_q \sum_{k=1}^{K} \mathbf{f}_{q,k} s_k + \tilde{\mathbf{n}}_q, \tag{2}$$

where λ_q represents the quantization impairment factor at the qth AP given by $\lambda_q = \sqrt{1 - \frac{\pi\sqrt{3}}{2} 2^{-2b_q}}$ with b_q being the resolution of the DACs at the qth AP. Moreover, $\tilde{\mathbf{n}}_q \sim \mathcal{CN}\left(\mathbf{0}_{N_t}, \sigma_{e,q}^2 \mathbf{I}_{N_t}\right)$ describes the quantization noise that is uncorrelated with $\mathbf{x}_q$, where $\sigma_{e,q}^2$ is the quantization noise variance expressed as $\sigma_{e,q}^2 = {\lambda_q}^2 \left(1 - \lambda_q^2\right)^2$ [4], [13].

The channel vector from the qth AP to the kth user, qth AP to the lth RIS, and lth RIS to the kth user is respectively denoted as $\mathbf{h}_{q,k} \in \mathbb{C}^{N_t \times 1}$, $\mathbf{G}_{q,l} \in \mathbb{C}^{M \times N_t}$, and $\mathbf{w}_{l,k} \in \mathbb{C}^{M \times 1}$. Moreover, we assume perfect channel state information[1], an ideal feedback link between the AP and the RIS controller, and unlimited capacity of the backhaul links from the APs to the CPU. The phase shift matrix of the lth RIS is given by $\mathbf{\Theta}_l = \text{diag}\left(\rho_1^l, ..., \rho_M^l\right) \in \mathbb{C}^{M \times M}$, where ρ_m^l denotes the frequency response of the mth element in the lth RIS with $\left|\rho_m^l\right| \leq 1$ [15]. The composite channel, composed of the direct and indirect links between the qth AP to the kth user, can be expressed by $\mathbf{z}_{q,k}^H = \mathbf{h}_{q,k}^H + \sum_{l=1}^{L} \mathbf{w}_{l,k}^H \mathbf{\Theta}_l \mathbf{G}_{q,l}$. Then, the received signal at the kth user is given by

$$y_k = \sum_{q=1}^{Q} \mathbf{z}_{q,k}^H \tilde{\mathbf{x}}_q + n_k, \tag{3a}$$

$$= \sum_{q=1}^{Q} \lambda_q \mathbf{z}_{q,k}^H \mathbf{f}_{q,c} s_c + \sum_{q=1}^{Q} \lambda_q \mathbf{z}_{q,k}^H \mathbf{f}_{q,k} s_k + \sum_{q=1}^{Q} \sum_{\substack{k'=1 \\ k' \neq k}}^{K} \lambda_q \mathbf{z}_{q,k}^H \mathbf{f}_{q,k'} s_{k'} + \sum_{q=1}^{Q} \mathbf{z}_{q,k}^H \tilde{\mathbf{n}}_q + n_k, \tag{3b}$$

where $n_k \in \mathcal{CN}\left(0, \sigma_k^2\right)$ represents the additive white Gaussian noise at the kth user. The signal-to-interference-and-noise-ratio (SINR) at the kth user to decode the common stream s_c is expressed as

$$\text{SINR}_k^c = \frac{\left| \sum_{q=1}^{Q} \lambda_q \mathbf{z}_{q,k}^H \mathbf{f}_{q,c} \right|^2}{\sum_{k'=1}^{K} \left| \sum_{q=1}^{Q} \lambda_q \mathbf{z}_{q,k}^H \mathbf{f}_{q,k'} \right|^2 + \sum_{q=1}^{Q} \mathbf{z}_{q,k}^H \tilde{\mathbf{n}}_q \tilde{\mathbf{n}}_q^H \mathbf{z}_{q,k} + \sigma_k^2}. \tag{4}$$

[1]We consider time-division duplex transmission in our system model, with our work focusing on downlink performance analysis. In practice, efficient channel estimation algorithms can be used by exploiting the sparsity and correlation of AP-RIS-user cascaded channels in mmWave RIS systems [14].

The SINR of private stream s_k at the kth user is given by

$$\mathrm{SINR}_k^p = \frac{\left|\sum_{q=1}^{Q} \lambda_q \mathbf{z}_{q,k}^H \mathbf{f}_{q,k}\right|^2}{\sum_{\substack{k'=1,\\ k'\neq k}}^{K} \left|\sum_{q=1}^{Q} \lambda_q \mathbf{z}_{q,k}^H \mathbf{f}_{q,k'}\right|^2 + \sum_{q=1}^{Q} \mathbf{z}_{q,k}^H \tilde{\mathbf{n}}_q \tilde{\mathbf{n}}_q^H \mathbf{z}_{q,k} + \sigma_k^2}. \tag{5}$$

The corresponding achievable rates at the kth user for the common stream s_c and private stream s_k are given by $\mathrm{R}_k^c = \log_2\left(1+\mathrm{SINR}_k^c\right)$ and $\mathrm{R}_k^p = \log_2\left(1+\mathrm{SINR}_k^p\right)$, respectively. At each user, the common stream is first decoded, and the SIC procedure is then used to decode its private stream. To ensure that the common stream can be decoded by all users, the transmission rate of the common stream, R_c, must not exceed the achievable rate for the common stream at the users, i.e. $R_c \leq \min\left\{R_1^c, ..., R_K^c\right\}$. Moreover, let c_k be the rate associate with the kth part of the super-common message, m_k^c, such that $R_c = \sum_{k=1}^{K} c_k$.

Our aim is to maximize the minimum rate at the users while satisfying the maximum available power at each AP. Hence, the max-min problem can be formulated as follows:

$$\max_{\mathbf{f}_{q,c},\mathbf{f}_{q,k},c_k,\rho_m^l} \ \min_k \ R_k^p + c_k \tag{6}$$

$$\text{s.t.} \ \sum_{i=1}^{K} c_i \leq R_k^c, \ \forall k \tag{6a}$$

$$\mathbf{f}_{q,c}^H \mathbf{f}_{q,c} + \sum_{k=1}^{K} \mathbf{f}_{q,k}^H \mathbf{f}_{q,k} \leq P_{q,\max}, \ \forall q \tag{6b}$$

$$c_k \geq 0, \ \forall k \tag{6c}$$

$$\left|\rho_m^l\right| \leq 1, \ \forall m, \forall l, \tag{6d}$$

where $P_{q,\max}$ is the maximum available power at the qth AP. Constraint (6a) ensures that the common stream can be decoded by all users. In the following section, we describe our proposed solution to the non-convex problem (6).

III. Problem Solution

The precoding design for the private precoding vectors is based on the maximum ratio transmission (MRT), which is widely used in CF-mMIMO systems [1]-[3], such that $\mathbf{f}_{q,k} = \alpha_{q,k}\tilde{\mathbf{f}}_{q,k}$, where $\alpha_{q,k}^2$ represents the transmit power and $\tilde{\mathbf{f}}_{q,k} = \frac{\mathbf{z}_{q,k}}{\|\mathbf{z}_{q,k}\|}$. Next, we include the auxiliary variable t and reformulate problem (6) as follows:

$$\max_{\mathbf{f}_{q,c},\mathbf{f}_{q,k},c_k,\rho_m^l} \ t \tag{7}$$

$$\text{s.t.} \ R_k^p + c_k \geq t, \ \forall k \tag{7a}$$

$$\text{(6a) - (6d).} \tag{7b}$$

Problem (7) is non-convex due to constraints (7a) and (6a). Therefore, we propose a AO-based algorithm to divide the Problem (7) into two subproblems. Initially, given the phase-shift matrices of RISs, we propose a SCA-based scheme to optimize the precoding vectors for the common stream, the power allocation variables, and the common rates. Subsequently, given the precoding vectors, we optimize the phase-shift matrices and the common rates.

A. Optimize $\mathbf{f}_{q,c}, \alpha_{q,k}, c_k$ for a given $\mathbf{\Theta}_l$

First, we stack the precoding vectors, the composite channels and variables related to the distortion of the quantization. These are given by the following expressions: $\mathbf{f}_c = \left[\mathbf{f}_{1,c}^T, ..., \mathbf{f}_{Q,c}^T\right]^T$, $\mathbf{f}_k = \left[\mathbf{f}_{1,k}^T, ..., \mathbf{f}_{Q,k}^T\right]^T$, $\boldsymbol{\alpha}_k = \left[\alpha_{1,k}, ..., \alpha_{Q,k}\right]^T$, $\mathbf{z}_k = \left[\mathbf{z}_{1,k}^T, ..., \mathbf{z}_{Q,k}^T\right]^T$, $\boldsymbol{\lambda} = \mathrm{diag}\left(\lambda_1, ..., \lambda_q\right)$, $\boldsymbol{\sigma}_e = \mathrm{diag}\left(\sigma_{e,1}^2, ..., \sigma_{e,Q}^2\right)$, and $\tilde{\mathbf{n}} = \left[\tilde{\mathbf{n}}_1^T, ..., \tilde{\mathbf{n}}_Q^T\right]^T$. Moreover, the term $\left|\sum_{q=1}^{Q} \lambda_q \mathbf{z}_{q,k}^H \mathbf{f}_{q,k'}\right|^2$ can be reformulated as $\left|\boldsymbol{\kappa}_{k,k'}^T \boldsymbol{\alpha}_k\right|^2$, where $\boldsymbol{\kappa}_{k,k'} = \left[\lambda_1 \mathbf{z}_{1,k}^H \tilde{\mathbf{f}}_{1,k'}, ..., \lambda_Q \mathbf{z}_{Q,k}^H \tilde{\mathbf{f}}_{Q,k'}\right]^T$.

Next, we reformulate the non-convex constraint (7a) by introducing the auxiliary variables J_k and v_k as follows:

$$\frac{\left|\boldsymbol{\kappa}_{k,k}^T \boldsymbol{\alpha}_k\right|^2}{J_k} \geq v_k - 1, \tag{8a}$$

$$\sum_{k'=1,k'\neq k}^{K} \left|\boldsymbol{\kappa}_{k,k'}^T \boldsymbol{\alpha}_{k'}\right|^2 + \mathbf{z}_k^H \left(\boldsymbol{\sigma}_e \otimes \mathbf{I}_{N_t}\right) \mathbf{z}_k + \sigma_k^2 \leq J_k, \tag{8b}$$

$$2^{t-c_k} \leq v_k. \tag{8c}$$

Similarly, by introducing the auxiliary variables $J_{c,k}$ and v_c, the non-convex constraint (6a) can be reformulate to

$$\frac{\left|\mathbf{z}_k^H \left(\boldsymbol{\lambda} \otimes \mathbf{I}_{N_t}\right) \mathbf{f}_c\right|^2}{J_{c,k}} \geq c_c - 1, \tag{9a}$$

$$\sum_{k'=1}^{K} \left|\boldsymbol{\kappa}_{k,k'}^T \boldsymbol{\alpha}_{k'}\right|^2 + \mathbf{z}_k^H \left(\boldsymbol{\sigma}_e \otimes \mathbf{I}_{N_t}\right) \mathbf{z}_k + \sigma_k^2 \leq J_{c,k}, \tag{9b}$$

$$2^{\sum_{k'=1}^{K} c_{k'}} \leq v_c. \tag{9c}$$

Next, we adopt the linear approximation from [9] to transform the non-convex constraint (8a), given the feasible values $\left\{\boldsymbol{\alpha}_k^{(n)}, J_k^{(n)}\right\}$ at the nth iteration, as follows:

$$\frac{2\Re\left(\left(\boldsymbol{\alpha}_k^{(n)}\right)^T \boldsymbol{\kappa}_{k,k} \boldsymbol{\kappa}_{k,k}^T \boldsymbol{\alpha}_k\right)}{J_k^{(n)}} - J_k \left(\frac{\left|\boldsymbol{\kappa}_{k,k}^T \boldsymbol{\alpha}_k^{(n)}\right|}{J_k^{(n)}}\right)^2 \geq v_k - 1. \tag{10}$$

Similarly, the non-convex constraint (9a), given the feasible values $\left\{\mathbf{f}_c^{(n)}, J_{c,k}^{(n)}\right\}$ at the nth iteration, is reformulated to

$$\frac{2\Re\left(\left(\mathbf{f}_c^{(n)}\right)^H \left(\boldsymbol{\lambda} \otimes \mathbf{I}_{N_t}\right)^H \mathbf{z}_k \mathbf{z}_k^H \left(\boldsymbol{\lambda} \otimes \mathbf{I}_{N_t}\right) \mathbf{f}_c\right)}{J_{c,k}^{(n)}} - J_{c,k} \left(\frac{\left|\mathbf{z}_k^H \left(\boldsymbol{\lambda} \otimes \mathbf{I}_{N_t}\right) \mathbf{f}_c^{(n)}\right|}{J_{c,k}^{(n)}}\right)^2 \geq v_c - 1. \tag{11}$$

Algorithm 1 Proposed solution for problem (7), given $\mathbf{\Theta}_l$

1: **inputs:** Counter n=0, initial points $\left\{\boldsymbol{\alpha}_k^{(n)}, J_k^{(n)}, \mathbf{f}_c^{(n)}, J_{c,k}^{(n)}\right\}, t^{(n)}$, and tolerance υ.
2: **while** $t^{(n)} - t^{(n-1)} > \upsilon$ **do**
3: Solve problem (12) with $\left\{\boldsymbol{\alpha}_k^{(n)}, J_k^{(n)}, \mathbf{f}_c^{(n)}, J_{c,k}^{(n)}\right\}$ and denote the solution as $\mathbf{f}_c^*, \nu_c^*, t^*, \{\boldsymbol{\alpha}_k^*, J_k^*, J_{c,k}^*, c_k^*, \nu_k^*\}$.
4: Update $\mathbf{f}_c^{(n+1)} \leftarrow \mathbf{f}_c^*$, $\boldsymbol{\alpha}_k^{(n+1)} \leftarrow \boldsymbol{\alpha}_k^*$.
5: Update $n \leftarrow n+1$.
6: **end while**
7: **outputs:** $\mathbf{f}_c^*, t^*, \{\boldsymbol{\alpha}_k^*, c_k^*\}$.

Algorithm 2 Proposed solution for problem (7), given $\mathbf{f}_{q,c}, \mathbf{f}_{q,k}$

1: **inputs:** Counter n=0, initial points $\left\{\mathbf{z}_k^{(n)}, S_k^{(n)}, S_{c,k}^{(n)}\right\}, t^{(n)}$, and tolerance value υ_R.
2: **while** $t^{(n)} - t^{(n-1)} > \upsilon_R$ **do**
3: Solve problem (15) with $\left\{\mathbf{z}_k^{(n)}, S_k^{(n)}, S_{c,k}^{(n)}\right\}$ and denote the solution as $\delta_c^*, t^*, \{\boldsymbol{\varphi}_l^*, c_k^*, S_k^*, S_{c,k}^*, \delta_k^*\}$.
4: Update $\boldsymbol{\varphi}_l^{(n+1)} \leftarrow \boldsymbol{\varphi}_l^*$, $S_{c,k}^{(n+1)} \leftarrow S_{c,k}^*$,$S_k^{(n+1)} \leftarrow S_k^*$, $t^{(n+1)} \leftarrow t^*$.
5: Update $n \leftarrow n+1$.
6: **end while**
7: **outputs:** $t^*, \{\boldsymbol{\varphi}_l^*, c_k^*\}$.

Hence, problem (7), given the phase-shift matrices of the RISs, can be reformulated using the feasible points $\left\{\boldsymbol{\alpha}_k^{(n)}, J_k^{(n)}, \mathbf{f}_c^{(n)}, J_{c,k}^{(n)}\right\}$ as follows:

$$\max_{\substack{\{\boldsymbol{\alpha}_k, v_k, c_k, J_k, J_{c,k}\}, \\ \mathbf{f}_c, t, v_c}} \quad t \tag{12}$$

$$\text{s.t. } \|\mathbf{f}_{q,c}\|^2 + \sum_{k=1}^{K} (\alpha_{q,k})^2 \le P_{q,\max}, \forall q \tag{12a}$$

$$c_k, t, v_c, v_k, J_k J_{c,k} \ge 0, \forall k \tag{12b}$$

$$\boldsymbol{\alpha}_k \ge \mathbf{0}_Q, \forall k, \tag{12c}$$

$$\text{(8b), (8c), (9b), (9c), (10), (11),} \tag{12d}$$

where $\alpha_{q,k} = \mathbf{a}_q \boldsymbol{\alpha}_k$ with $\mathbf{a}_q = \left[\mathbf{0}^{1\times(q-1)}, 1, \mathbf{0}^{1\times(Q-q)}\right]$, and $\mathbf{p}_{q,c} = \mathbf{A}_q \mathbf{p}_c$ with $\mathbf{A}_q = \left[\mathbf{0}^{N_t\times(q-1)N_t}, \mathbf{I}_{N_t}, \mathbf{0}^{N_t\times(Q-q)N_t}\right]$. Problem (12) is convex, and the CVX toolbox [16] can be used to obtain the solution. Problem (12) consists of one variable of size QN and $\varsigma = 2 + K(4+Q)$ real variables having a computational complexity of $\mathcal{O}\left((QN+\varsigma)^{3.5}\right)$ [10]. Algorithm 1 describes the overall procedure for solving the problem (7) given $\mathbf{\Theta}_l$.

B. Optimize $\mathbf{\Theta}_l, c_k$ for a given $\mathbf{f}_{q,c}$ and $\mathbf{f}_{q,k}$

By introducing $\boldsymbol{\varphi}_l = \left[\rho_1^l, ..., \rho_M^l\right]^T$, the composite channel is defined as $\mathbf{z}_{q,k}^H = \mathbf{h}_{q,k}^H + \sum_{l=1}^{L} \boldsymbol{\varphi}_l^T \text{diag}\left(\mathbf{w}_{l,k}^H\right)\mathbf{G}_{q,l}$. Next, for constraint (7a), we first introduce the auxiliary variables S_k and δ_k. By following a similar procedure used for (8a)-(8c) and applying the linear approximation given the feasible points $\left\{\mathbf{z}_k^{(n)}, S_k^{(n)}\right\}$, where $\mathbf{z}_k^{(n)} = [(\mathbf{z}_{1,k}^{(n)})^T, ..., (\mathbf{z}_{Q,k}^{(n)})^T]^T$ with $\left(\mathbf{z}_{q,k}^{(n)}\right)^H = \mathbf{h}_{q,k}^H + \sum_{l=1}^{L} \left(\boldsymbol{\varphi}_l^{(n)}\right)^T \text{diag}\left(\mathbf{w}_{l,k}^H\right)\mathbf{G}_{q,l}$, constraint (7a) can be reformulated as follows:

$$\frac{2\Re\left(\mathbf{f}_k^H(\boldsymbol{\lambda}\otimes\mathbf{I}_{N_t})^H \mathbf{z}_k^{(n)} \mathbf{z}_k^H (\boldsymbol{\lambda}\otimes\mathbf{I}_{N_t})\mathbf{f}_k\right)}{S_k^{(n)}} - S_k\left(\frac{\left|\left(\mathbf{z}_k^{(n)}\right)^H (\boldsymbol{\lambda}\otimes\mathbf{I}_{N_t})\mathbf{f}_k\right|}{S_k^{(n)}}\right)^2 \ge \delta_k - 1. \tag{13a}$$

$$\sum_{k'=1,k'\neq k}^{K} \left|\mathbf{z}_k^H (\boldsymbol{\lambda}\otimes\mathbf{I}_{N_t})\mathbf{f}_{k'}\right|^2 + \mathbf{z}_k^H (\boldsymbol{\sigma}_e\otimes\mathbf{I}_{N_t})\mathbf{z}_k + \sigma_k^2 \le S_k, \tag{13b}$$

$$2^{t-c_k} \le \delta_k. \tag{13c}$$

Similarly, for constraint (6a), we first include the auxiliary variables $S_{c,k}$ and δ_c. Then, by following a similar procedure used for (9a)-(9c) along with the linear approximation given the feasible points $\left\{\mathbf{z}_k^{(n)}, S_{c,k}^{(n)}\right\}$, constraint (6a) can be reformulated as follows:

$$\frac{2\Re\left(\mathbf{f}_c^H(\boldsymbol{\lambda}\otimes\mathbf{I}_{N_t})^H \mathbf{z}_k^{(n)} \mathbf{z}_k^H (\boldsymbol{\lambda}\otimes\mathbf{I}_{N_t})\mathbf{f}_c\right)}{S_{c,k}^{(n)}} - S_{c,k}\left(\frac{\left|\left(\mathbf{z}_k^{(n)}\right)^H (\boldsymbol{\lambda}\otimes\mathbf{I}_{N_t})\mathbf{f}_c\right|}{S_{c,k}^{(n)}}\right)^2 \ge \delta_c - 1. \tag{14a}$$

$$\sum_{k'=1}^{K} \left|\mathbf{z}_k^H (\boldsymbol{\lambda}\otimes\mathbf{I}_{N_t})\mathbf{f}_{k'}\right|^2 + \mathbf{z}_k^H (\boldsymbol{\sigma}_e\otimes\mathbf{I}_{N_t})\mathbf{z}_k + \sigma_k^2 \le S_{c,k}, \tag{14b}$$

$$2^{\sum_{k'=1}^{K} c_{k'}} \le \delta_c. \tag{14c}$$

Finally, problem (7), given the precoding vectors and the feasible points $\left\{\mathbf{z}_{q,k}^{(n)}, S_k^{(n)}, S_{c,k}^{(n)}\right\}$, at the nth iteration can be expressed as

$$\max_{\substack{\{c_k, \delta_k, S_k, S_{c,k}\}, \\ \boldsymbol{\varphi}_l, t, \delta_c}} \quad t \tag{15}$$

$$\text{s.t. } \left|\rho_m^l\right| \le 1, \forall m, \forall l \tag{15a}$$

$$t, c_k, S_k, S_{c,k}, \delta_k, \delta_c \ge 0, \forall k \tag{15b}$$

$$\text{(13a)-(13c), (14a)-(14c).} \tag{15c}$$

Problem (15) is convex and can be solved by the CVX toolbox [16]. Problem (15) consists of L variables of size M and $\varsigma_R = 2+4K$ real variables having a computational complexity of $\mathcal{O}\left((LM+\varsigma_R)^{3.5}\right)$ [10]. Algorithm 2 describes the overall procedure for solving the problem (7) given $\mathbf{f}_{q,c}, \mathbf{f}_{q,k}$. Finally, Algorithm 1 and Algorithm 2 are used in an iterative manner to optimize the precoding vectors, common rates, and phase shift matrices until convergence.

IV. Simulation Results

In our simulations, the Saleh-Valenzuela channel model was adopted, and the operating frequency was set to 30 GHz. The APs and RISs are equipped with uniform planar arrays (UPAs).

We consider a scenario in a coverage area of 40m×40m with 4 users, setting the noise power at -90 dBm. Subsequently, the channels can be expressed as follows [6]:

$$\mathbf{h}_{q,k} = \sqrt{\frac{N_t}{C^h_{q,k}}} \sum_{c=1}^{C^h_{q,k}} \omega^h_{q,k,c} \mathbf{a}_T \left(\gamma^{h,T}_{q,k,c}, \eta^{h,T}_{q,k,c}\right), \tag{16a}$$

$$\mathbf{G}_{q,l} = \sqrt{\frac{N_t M}{C^G_{q,l}}} \sum_{c=1}^{C^G_{q,l}} \omega^G_{q,l,c} \mathbf{a}_R \left(\gamma^{G,R}_{q,l,c}, \eta^{G,R}_{q,l,c}\right) \mathbf{a}^H_T \left(\gamma^{G,T}_{q,l,c}, \eta^{G,T}_{q,l,c}\right), \tag{16b}$$

$$\mathbf{w}_{l,k} = \sqrt{\frac{M}{C^w_{l,k}}} \sum_{c=1}^{C^w_{l,k}} \omega^w_{l,k,c} \mathbf{a}_T \left(\gamma^{w,T}_{l,k,c}, \eta^{w,T}_{l,k,c}\right), \tag{16c}$$

where $C^h_{q,k}$ represents the number of paths for the channel $\mathbf{h}_{q,k}$, and $\gamma^{h,T}_{q,k,c}$, $\eta^{h,T}_{q,k,c}$ denote the azimuth and elevation angles in th cth path for the channel $\mathbf{h}_{q,k}$. The number of paths is set to $C^h_{q,k} = 2$, $C^G_{q,l} = 3$, $C^w_{l,k} = 3$, and the path angles are randomly selected within $[0, 2\pi]$. The superscripts R and T on the path angles stand for the arrival and departure, respectively. $\omega^h_{q,k,c}$ denotes the complex channel gain, which is given by $\omega_i \sim CN\left(0, 10^{-0.1P_L(d_i)}\right)$. Here, $P_L(d_i)$ represents the path loss at a distance d_i and is given by [6, Eq. (46)], where we assume that the AP-RIS and RIS-User links are LoS, while the AP-user direct links are NLoS. Moreover, $\mathbf{a}_{T(R)}(\gamma, \eta)$ stands for the array response vector for half-wavelength UPA having $N = N_y \times N_z$ elements and is given by [6, Eq. (9)]. For comparative purposes, we consider the traditional scheme that does not incorporate RSMA, which we denote as 'NoRS', and also the method that does not involve the deployment of RISs. The NoRS method, which does not divide the message of users, can be formulated based on Problem (6) by eliminating the precoding vector for the common stream, the common rate variables, and constraints (6a) and (6c).

First, we consider a random deployment of APs, RISs, and users in the coverage area, where 6 RISs are deployed, each having $M = 64$ reflective elements each. The total power allocated to all APs in the system is 2W, and the total number of antennas is 72. Therefore, when the number of APs increases, the total power and the total number of antennas of the system are divided equally among the deployed APs, ensuring a fair performance comparison. Fig. 2 presents the max-min rate at the users versus the number of APs, Q, considering two different bit resolutions, $b_q = b = \{1, 4\}$, at the DACs. We observe that increasing the number of APs leads to an increase in the max-min rate of the users because each AP can be located closer to the users and RISs. In this deployment, the performance gain due to RISs is not very significant due to the random deployment of RISs, which does not guarantee that an RIS will be located close to the intended users. Moreover, we can see the improvement in the rate due to RSMA over the traditional scheme because RSMA can manage user interference by using the common stream and the SIC layer at the users. Furthermore, we can observe a lower max-min rate when reducing the bit resolution at the DACs due to the increase in the distortion factor and quantization noise.

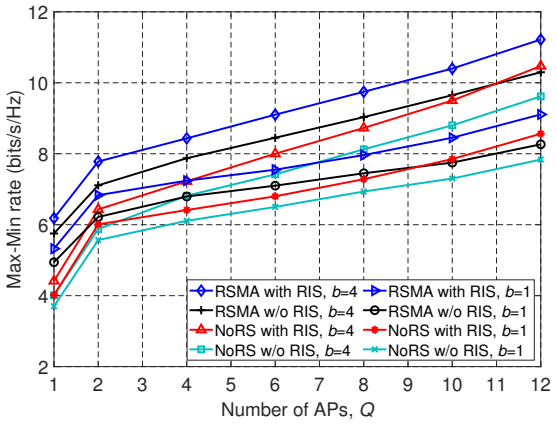

Fig. 2. Max-min rate versus the number of APs in a random deployment.

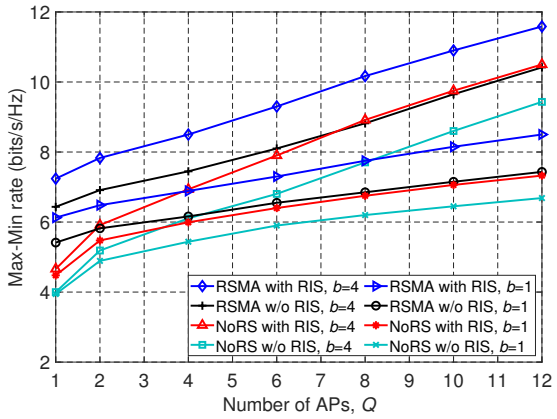

Fig. 3. Max-min rate versus the number of APs considering RISs located near the users.

However, it is a known fact that reducing the bit resolution can lead to enhanced energy efficiency in practice.

Next, we investigate the RISs in a hotspot deployment where the RISs are located closer to the users. Specifically, we consider that APs are randomly deployed in the coverage area, while the RISs and users are randomly deployed within a circular area with a radius of 8m located in the center of the coverage area. We consider 4 RISs with 64 reflective elements each, maintaining the same total power and total number of antennas as in the previous scenario. From Fig. 3, we can observe an improvement in the max-min rate by deploying RISs in the system. This improvement can be observed for both schemes, RSMA and without RSMA, when comparing the cases with RIS and without RIS. We can observe that a carefully planned RIS deployment is required to significantly increase the minimum achievable rate of the users, in contrast to the random deployment analyzed in Fig. 2, which includes two additional RISs compared to the scenario in Fig. 3.

The convergence behavior of the proposed scheme is presented in Fig. 4, where we consider 3 RISs with 144 reflecting elements, 4 APs with $N_t = 9$ antennas, and $P_{\max} = 100$ mW

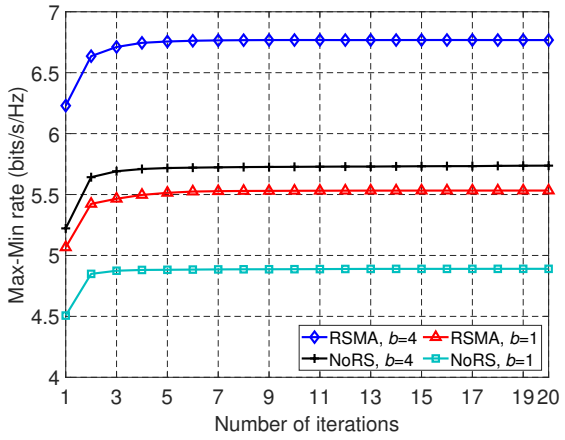

Fig. 4. Convergence behavior.

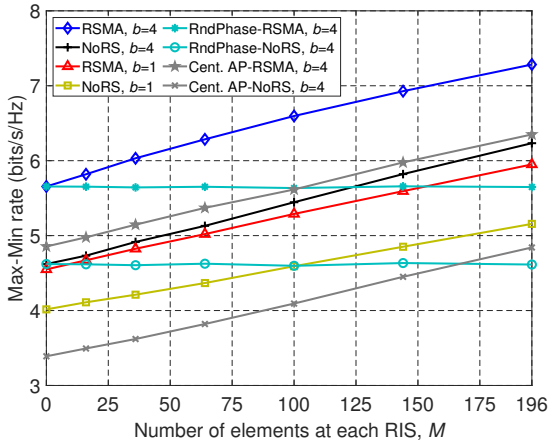

Fig. 5. Max-min rate versus the number of elements in each RIS.

per AP. We can observe that the proposed scheme achieves convergence within 8 iterations. Finally, Fig. 5 presents the max-min rate for the users versus the number of elements in each RIS by considering 3 RISs, 4 APs with $N_t = 9$, and $P_{\max} = 100$ mW per AP. We consider the baseline schemes of "RndPhase" which involves randomly selecting the phase shifts of RISs along with Algorithm 1, and the "Cent. AP," which involves serving all users with a centralized AP equipped with 36 antennas and a transmission power of 400 mW. We can observe a considerable improvement in the max-min rate for all the compared schemes as the number of RIS elements increases, due to the greater degree of freedom provided by the RISs. Moreover, the proposed approach outperforms other baseline schemes, highlighting the need to optimize the reflection coefficients and the improvement of distributed APs instead of the centralized approach.

V. Conclusion

In this paper, we propose an RSMA-based framework in a multi-RIS-assisted CF-mMIMO system, considering low-resolution DACs at the APs. We aim to maximize the minimum rate among the users by optimizing the common and private precoding vectors, common rate variables, and the reflection coefficients of RISs. The proposed solution relies on the AO algorithm and SCA methods. Simulation results for a mmWave system demonstrate that our proposed RSMA-based approach outperforms the traditional scheme that does not incorporate RSMA. Furthermore, we investigate various deployment strategies for RISs, illustrating that the minimum rate for the users can be significantly enhanced by strategically planning the deployment of RISs within the system.

References

[1] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.

[2] H. Q. Ngo, L.-N. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the Total Energy Efficiency of Cell-Free Massive MIMO," *IEEE Trans. Green Commun. Netw.* , vol. 2, no. 1, pp. 25–39, Mar. 2018.

[3] H. D. Tuan, A. A. Nasir, H. Q. Ngo, E. Dutkiewicz, and H. V. Poor, "Scalable User Rate and Energy-Efficiency Optimization in Cell-Free Massive MIMO," *IEEE Trans. Commun.*, vol. 70, no. 9, pp. 6050–6065, Sep. 2022.

[4] A. Kaushik, E. Vlachos, C. Tsinos, J. Thompson, and S. Chatzinotas, "Joint Bit Allocation and Hybrid Beamforming Optimization for Energy Efficient Millimeter Wave MIMO Systems," *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 1, pp. 119–132, Mar. 2021.

[5] I. Kim, M. Bennis, and J. Choi, "Cell-Free mmWave Massive MIMO Systems With Low-Capacity Fronthaul Links and Low-Resolution ADC/DACs," *IEEE Trans. Veh. Technol.*, vol. 71, no. 10, pp. 10512–10526, Oct. 2022.

[6] B. Ning, P. Wang, L. Li, Z. Chen, and J. Fang, "Multi-IRS-Aided Multi-User MIMO in mmWave/THz Communications: A Space-Orthogonal Scheme," *IEEE Trans. Commun.*, vol. 70, no. 12, pp. 8138–8152, Dec. 2022.

[7] Z. Zhang and L. Dai, "A Joint Precoding Framework for Wideband Reconfigurable Intelligent Surface-Aided Cell-Free Network," *IEEE Trans. Signal Process.*, vol. 69, pp. 4085–4101, 2021.

[8] K. Wang et al., "Reconfigurable Intelligent Surfaces Aided Energy Efficiency Maximization in Cell-Free Networks," *IEEE Wireless Commun. Lett.*, vol. 13, no. 6, pp. 1596–1600, Jun. 2024.

[9] Y. Mao, B. Clerckx, and V. O. K. Li, "Rate-Splitting for Multi-Antenna Non-Orthogonal Unicast and Multicast Transmission: Spectral and Energy Efficiency Analysis," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8754–8770, Dec. 2019.

[10] M. R. Camana, C. E. Garcia, and I. Koo, "Rate-Splitting Multiple Access in a MISO SWIPT System Assisted by an Intelligent Reflecting Surface," *IEEE Trans. Green Commun. Netw.*, vol. 6, no. 4, pp. 2084–2099, Dec. 2022.

[11] W. Wang, L. Li, G. Deng, and J. Li, "A Joint Multiservice Transmission Scheme for RSMA-Aided Cell-Free mMIMO System," *IEEE Commun. Lett.*, vol. 27, no. 2, pp. 591–594, Feb. 2023.

[12] A. Mishra, Y. Mao, L. Sanguinetti, and B. Clerckx, "Rate-Splitting Assisted Massive Machine-Type Communications in Cell-Free Massive MIMO", *IEEE Commun. Lett.*, vol. 26, no. 6, pp. 1358–1362, Jun. 2022.

[13] O. Dizdar, A. Kaushik, B. Clerckx, and C. Masouros, "Energy Efficient Dual-Functional Radar-Communication: Rate-Splitting Multiple Access, Low-Resolution DACs, and RF Chain Selection," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 986–1006, 2022.

[14] G. Zhou, C. Pan, H. Ren, P. Popovski, and A. L. Swindlehurst, "Channel Estimation for RIS-Aided Multiuser Millimeter-Wave Systems," *IEEE Transactions on Signal Processing*, vol. 70, pp. 1478–1492, 2022

[15] W. Yan et al., "Beamforming Analysis and Design for Wideband THz Reconfigurable Intelligent Surface Communications," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 8, pp. 2306–2320, Aug. 2023.

[16] M. C. Grant and S. P. Boyd. "CVX: MATLAB Software for Disciplined Convex Programming, Version 2.2." Jul. 2024. [Online]. Available: http://cvxr.com/cvx