RSMA for Reconfigurable Holographic Surface-aided Downlink Near-Field Communication

Sandeep Singh[†], Keshav Singh[†], Sandeep Kumar Singh[‡], Aryan Kaushik[⋆], and Trung Q. Duong[§]
[†]Institute of Communications Engineering, National Sun Yat-sen University, Kaohsiung 80424, Taiwan
[‡]Department of ECE, Motilal Nehru National Institute of Technology Allahabad, India
[§]Dept. of Computing and Mathematics, Manchester Metropolitan University, Manchester M15 6BX, UK
[§]Faculty of Engineering and Applied Science, Memorial University, NL A1C 5S7, Canada
E-mail: m113620004@student.nsysu.edu.tw, ksingh1980@ieee.org, sksingh@mnnit.ac.in, a.kaushik@ieee.org, tduong@mun.ca

***Abstract*—This work integrates rate-splitting multiple access (RSMA) in a reconfigurable intelligent surface (RIS)-aided holographic multi-input-multi-output (MIMO) downlink wireless communication system under near-field coverage. To maximize the sum rate of the system, we jointly optimize the transmit beamforming at the Base station (BS) and passive beamforming at the RIS. We suggest an iterative solution based on alternating optimization (AO) that efficiently solves the joint optimization problem utilizing analytical tools such as the successive convex approximation. Finally, the numerical results are shown to confirm the proposed algorithm's convergence and efficacy. Furthermore, We also highlight the impact of reflecting elements, minimum rate constraint, maximum transmit power at the BS, and number of downlink users. The dominance of RSMA over conventional multiple-access technologies such as non-orthogonal multiple access (NOMA) is also demonstrated.**

***Index Terms*—Reconfigurable intelligent surface (RIS), holographic multi-input-multi-output (MIMO), sum-rate maximization, rate-splitting multiple access (RSMA), near-field.**

I. Introduction

NEXT-generation wireless networks are envisioned to provide high-speed data transmission to cope with the ever-increasing consumer demands. Consequently, the power demand also increases, which has a detrimental effect on the environment [1]. This has led to a dramatic shift from prioritizing spectral efficiency towards energy efficiency [1], [2]. Furthermore, with the recent advancement in technological development, both energy and spectral efficiency are improving [2]. Moreover, multifold improvements are enabled by the introduction of holographic multi-input multi-output (MIMO) technologies [3]. A holographic antenna is a two-dimensional (2D) antenna composed of numerous micro-antennas fabricated on a planar surface [3], [4]. These planar antennas also offer significant beamforming gains [5], resulting from their widened electromagnetic aperture with an almost infinite number of antenna elements. They are also classified into active and passive types, where one amplifies the incident signal while the other does not. Advocating the immense advantages of holographic surfaces, their impact on reconfigurable reflecting surfaces (RIS) is an interesting area of research that is still unexplored to a larger extent. The adaptation of RIS as holographic MIMO in sixth-generation wireless networks has led to a significant performance enhancement [3]. This improvement was achieved by incorporating metamaterials and subwavelength-sized antennas, which imitated a continuous aperture [3]. The dense arrangement also enables the control of the electromagnetic (EM) environment with deterministic ability envisaged through channel hardening [6].

Note that the above-discussed articles presented their studies under the consideration of far-field assumptions [7]. However, RIS is a technological advancement that has a wide spread of applications in wireless networks. For example, when RIS is operated at millimeter (mm)-wave frequencies, they exhibit a large aperture [8]. In the scenario of large aperture arrays, the adaptation of near-field assumptions and the concept of spherical waves is more suitable, which can reliably distinguish users by considering both distance and angle [9]. Compared to far-field communication, near-field communication provides higher spectral efficiency and a higher spatial degree of freedom. Further, depending on the distance (Rayleigh distance), deploying an RIS to provide wireless services within a nearby region is likely to cause communication to occur in the near field. A prominent characteristic of the near-field region is the dominance of the distance-aware spherical wavefront, which allows beamfocusing to deliver wireless energy to a spatial location configured with three-dimensional spherical polar co-ordinate [10]. For example, the feasibility of RIS-aided near-field communication was verified in [11], where closed-form expressions for near-field and far-field boundaries were derived. In [12], a RIS-aided near-field communication scenario was investigated, where the sum-rate maximization problem was addressed using the alternating optimization (AO) algorithm. The use of RIS-aided holographic MIMO modifies the channel gain to boost coverage and improve system performance, particularly in challenging channel conditions [13], [14].

Notably, as compared to the non-orthogonal multiple access (NOMA) scheme, the rate-splitting multiple access (RSMA) technique has been acknowledged as a more efficient, reliable, and strong multiple access technique for effective multi-user communication [15], [16]. The adaptation of RSMA in near-field-driven wireless systems is found to perform better than conventional multiple-access schemes [17]. Nonetheless, the continuous nature of holographic MIMO poses considerable

difficulties for array signal processing, rendering conventional beamforming techniques designed for traditional MIMO inefficient. Thus, there is a strong demand to address the unavoidable difficulties and challenges posed by holographic MIMO in order to have efficient wireless communication. To the best of our knowledge, this is the first study that investigates the use case of RSMA for RIS-aided holographic MIMO downlink wireless communication under the near-field framework.

In particular, this work investigates the integration of RSMA in RIS-aided holographic MIMO under a near-field-driven wireless communication system. A sum rate maximization problem is formulated that jointly optimizes key system resources, namely 1) the transmit beamforming vectors, 2) the common rate allocation vector, and 3) the RIS phase matrix. Due to the non-convex nature of the problem, it is divided into two subproblems that are iteratively solved. In order to tackle the non-convex nature of the formulated problem, we propose an alternating optimization (AO)-based iterative algorithm that effectively solves the joint optimization problem using analytical tools like successive convex approximation technique. Numerical analysis and simulations substantiate the efficiency and convergence of the proposed algorithm. It shows that the use of RSMA for RIS-aided holographic MIMO under a downlink near-field system attains a higher sum rate compared to the NOMA. Furthermore, we also demonstrate the impact of key system parameters, such as RIS elements, minimum quality of service (QOS) constraint corresponding to each user, and maximum transmit power at the base station (BS).

Notations: $\mathrm{Re}(\cdot)$ represents the real component. $||\cdot||$, $||\cdot||_F$, $\mathrm{Tr}(\cdot)$ $(\cdot)^T, (\cdot)^*$, and $(\cdot)^H$ represent the Euclidean norm, F-norm, trace, transpose, conjugate, and Hermitian conjugate, respectively, of the respective matrix/vector. $\lambda_{max}(\cdot)$ and $\mathbf{u}_{max}(\cdot)$ denote the biggest eigenvalue and the associated eigenvector. $\mathrm{diag}(\cdot)$ represents the diagonal matrix of the respective vector. $(\cdot)_{[i,j]}$ is the ith row and jth column of the respective matrix.

II. System Model

We consider a holographic RIS-aided muti-user downlink communication system consisting of a BS, a RIS and D downlink users $(\mathrm{DL}_1, \mathrm{DL}_2, \cdots, \mathrm{DL}_D)$, as shown in Fig. 1. The BS is equipped with M transmit antennas, while $\{\mathrm{DL}_d\}$ is equipped with a single receive antenna. Additionally, RIS has $K = A \times C$ passive elements positioned on the Y-Z plane, with its central point aligned at the origin of the coordinate system. Further, it is assumed that the direct link between BS and users is blocked due to unavoidable blockages/obstacles. Note that as stated earlier in the previous section, due to the smaller distance among the nodes, near-field driven channel assumptions are considered for the analysis.

A. Channel Model

1) Channel between $\{DL_d\}$ and RIS: When all the users are located in the near-field region, the transmitted signal should be modeled as a spherical wave and cannot be approximated

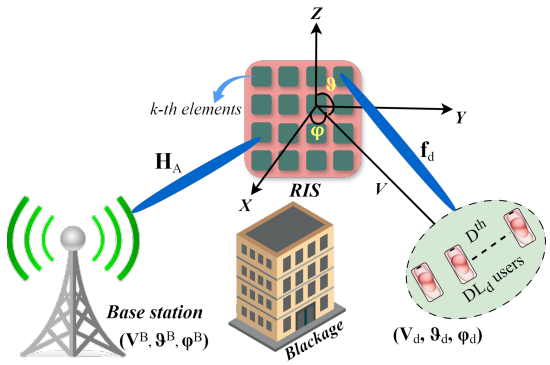

Fig. 1: System model.

as a plane wave for the RIS. For example, the channel between $\{\mathrm{DL}_d\}$ and the RIS is represented as [18]

$$\mathbf{f}_d(\mathbf{t}_d) = \frac{\sqrt{G}\lambda}{4\pi V_d} \exp\left(-j\frac{2\pi}{\lambda}V_d\right)\mathbf{a}_d, \tag{1}$$

where

$$\mathbf{a}_d = \exp\left(-j\frac{2\pi}{\lambda}\right)\Bigg\{-\sin\theta_d(A\mathfrak{d}_A\cos\varphi_d + C\mathfrak{d}_C\sin\varphi_d) + \frac{A^2\mathfrak{d}_A^2 + C^2\mathfrak{d}_C^2}{2V_d}\Bigg\}, \tag{2}$$

$\mathbf{t}_d = [V_d, \theta_d, \varphi_d]^T$ denotes the location of DL_d, with V_d, θ_d, φ_d being the range, elevation angle, and azimuth angle in the spherical coordinate system, respectively. G accounts for the radiation pattern and scale of RIS. The range of near field region, represented by Fresnel approximation to Rayleigh distance $\frac{2(\mathfrak{D}_A^2\mathfrak{D}_C^2)}{\lambda}$ large; where $\mathfrak{D}_A$ and $\mathfrak{D}_C$ are the aperture length along the Y-Z axis, $\mathfrak{d}_A = \frac{\mathfrak{D}_A}{A-1}$ and $\mathfrak{d}_C = \frac{\mathfrak{D}_C}{C-1}$ represents inter-element separation, while λ is the mm-Wave operating frequency [18], [19]. Similarly, the channel between BS - RIS $\mathbf{H}_A$ is also modeled.

B. System Model

In order to communicate with all D users, using the basic principle of RSMA, the BS will split each user's message into a common part and a private part. The information bits intended for the common and private sections of the d-th the BS specifically encodes the user's message into common and private symbols, expressed as x_c and $x_d, \forall d \in \mathcal{D}$, respectively. Note that x_c is obtained after combining the splitted common part of each user, while x_d corresponds to the splitted private part of the respective user. The BS prepares the combined symbol for transmission to each user, which is mathematically written as

$$\mathbf{x} = \underbrace{\mathbf{w}_c x_c}_{\text{Common signal}} + \underbrace{\sum_{d=1}^{D}\mathbf{w}_d x_d}_{\text{Private signal}}, \tag{3}$$

where $\mathbf{w}_c \in \mathbb{C}^{M\times 1}$ and $\mathbf{w}_d \in \mathbb{C}^{M\times 1}$ are the linear beamforming vectors for the common and private symbols, respectively, and $\mathbb{E}\{|x_c|^2\} = 1$ and $\mathbb{E}\{|x_d|^2\} = 1$. Further, we denote

$$\mathbf{h}_d = \mathbf{H}_A^H \mathbf{\Psi}\mathbf{f}_d(\mathbf{t}_d), \quad \forall d \in \mathcal{D}. \tag{4}$$

Here, in (4) the phase shift matrix at RIS is represented by $\mathbf{\Psi} \triangleq \text{diag}(\boldsymbol{\psi}) \in \mathbb{C}^{K\times K}$, $\boldsymbol{\psi} = [\exp(j\psi_1), \cdots, \exp(j\psi_K)]^T$. The phase shift of i^{th} RIS elements to the received signal is denoted by ψ_i. The signal received at DL$_d$ is given by

$$y_d = \mathbf{h}_d^H \mathbf{w}_c x_c + \sum_{d=1}^{D} \mathbf{h}_d^H \mathbf{w}_d x_d + n_d, \quad \forall d \in \mathcal{D}, \tag{5}$$

where n_d is the additive white Gaussian noise (AWGN), which is represented as a complex Gaussian random variable with a variance of σ_d^2 and a mean of zero, i.e., $n_d \sim \mathcal{CN}(0, \sigma_d^2)$.

Additionally, each user interprets the private symbol as interference while decoding the common symbol. As a result, the signal-to-interference-plus-noise ratio (SINR) for the common symbol at DL$_d$ can be written as

$$\gamma_{c,d} = \frac{|\mathbf{h}_d^H \mathbf{w}_c|^2}{\sum\limits_{i=1}^{D} |\mathbf{h}_d^H \mathbf{w}_i|^2 + \sigma_d^2}, \quad \forall d \in \mathcal{D}. \tag{6}$$

Thereafter, the common message is removed from the received signal using one-step SIC. Thereafter, each user decodes its own private symbol, considering the private symbols of all other users to be interference. Thus, the SINR for the private signal symbol received by the d-th user can be expressed by

$$\gamma_{p,d} = \frac{|\mathbf{h}_d^H \mathbf{w}_d|^2}{\sum\limits_{i=1, i\neq d}^{D} |\mathbf{h}_d^H \mathbf{w}_i|^2 + \sigma_d^2}, \quad \forall d \in \mathcal{D}. \tag{7}$$

The approximate achievable rate for both the common and private signal symbols at the d-th user can be computed as f

$$R_{c,d} = \log_2(1 + \gamma_{c,d}), \quad \forall d \in \mathcal{D}, \tag{8}$$

$$R_{p,d} = \log_2(1 + \gamma_{p,d}), \quad \forall d \in \mathcal{D}. \tag{9}$$

Note that the achievable rate of the common signal must not exceed the channel capacity corresponding to the common symbol. In particular, $\sum_{d=1}^{D} \varkappa_{c,d} \leq R_{c,d}, \forall d \in \mathcal{D}$, where $\varkappa_{c,d}$ is the the d-th user portion of the common rate. Therefore, the total achievable sum-rate of the considered RSMA under a downlink near-field system can be written as $R_{\text{Total}} = \sum_{d=1}^{D} (\varkappa_{c,d} + R_{p,d})$.

III. Problem Formulation

A. Problem Formulation

The primary objective of this work is to maximize the sum rate of the proposed RIS-aided holographic MIMO downlink near-field RSMA system by jointly designing the beamforming matrix $\mathbf{W} = [\mathbf{w}_c, \mathbf{w}_1, \cdots, \mathbf{w}_D]$ at the BS and phase-shift $\boldsymbol{\psi}$ at the RIS of common and private messages. Consequently, the optimization problem is formulated as

$$\mathcal{P}0 \quad \max_{\varkappa_c, \mathbf{W}, \boldsymbol{\psi}} \quad \sum\nolimits_{d=1}^{D} (\varkappa_{c,d} + R_{p,d}) \tag{10a}$$

$$\text{s.t.} \quad \sum\nolimits_{i=1}^{D} \varkappa_{c,i} \leq \log_2(1 + \gamma_{c,d}), \quad \forall d \in \mathcal{D}, \tag{10b}$$

$$\varkappa_{c,d} + R_{p,d} \geq R_d^{min}, \quad \forall d \in \mathcal{D}, \tag{10c}$$

$$R_{p,d} \leq \log_2(1 + \gamma_{p,d}), \quad \forall d \in \mathcal{D}, \tag{10d}$$

$$\text{Tr}\left(\mathbf{W}^H \mathbf{W}\right) \leq P_{max}, \tag{10e}$$

$$|\psi_k| = 1, \quad k \in \mathcal{K}, \tag{10f}$$

$$\varkappa_{c,d} \geq 0, \quad \forall d \in D, \tag{10g}$$

where $\varkappa_c = \{\varkappa_{c,d}\}$, (10b) corresponds to the common rate constraint, (10c) is QoS constraint which ensures the minimum required rate, (10e) is the BS transmit power constraint, and (10f) is the unit modulus constraint. Next, we discuss the AO-based solution to tackle the coupling of variables and the non-convex nature of the objective function.

B. Design Transmit Beamforming at BS

Note that using AO approach, the sub-optimal value of $\mathbf{W}$ can be obtained under the fixed phase shift at the RIS, i.e., $\boldsymbol{\psi} = \bar{\boldsymbol{\psi}}$. Consequently, the problem $\mathcal{P}0$ can be reformulated by incorporating the auxiliary variables $\boldsymbol{\Upsilon}_p, \boldsymbol{\beta}_p$, and $\boldsymbol{\chi}_c$

$$\mathcal{P}1 \quad \max_{\mathbf{W}, \varkappa_c, \boldsymbol{\Upsilon}_p, \boldsymbol{\beta}_p, \boldsymbol{\chi}_c} \quad \sum\nolimits_{d=1}^{D} (\varkappa_{c,d} + \Upsilon_{p,d}) \tag{11a}$$

$$\text{s.t.} \quad \sum\nolimits_{d=1}^{D} \varkappa_{c,d} \leq \log_2(1 + \chi_{c,d}), \quad \forall d \in \mathcal{D}, \tag{11b}$$

$$\varkappa_{c,d} + \Upsilon_{p,d} \geq R_d^{min}, \quad \forall d \in \mathcal{D}, \tag{11c}$$

$$\Upsilon_{p,d} \leq \log_2(1 + \beta_{p,d}), \quad \forall d \in \mathcal{D}, \tag{11d}$$

$$(6), (7), (10e), (10g),$$

where $\boldsymbol{\Upsilon}_p = \left[\Upsilon_{p,1}, \cdots, \Upsilon_{p,D}\right]$, $\boldsymbol{\beta}_p = \left[\beta_{p,1}, \cdots, \beta_{p,D}\right]$, $\boldsymbol{\chi}_c = \left[\chi_{c,1}, \cdots, \chi_{c,D}\right]$. $\overline{\mathbf{h}}_d$ is substituted as $\overline{\mathbf{h}}_d \triangleq \mathbf{h}_d|_{\boldsymbol{\psi}=\bar{\boldsymbol{\psi}}}$. Note that the non-concave objective function and the non-convexity of constraint (6) and (7) prevent a closed-form solution for the $\mathcal{P}1$. Therefore, in order to iteratively approximate the problem in $\mathcal{P}1$ into a similar convex form, we use the SCA approach. Furthermore, the terms in (6) and (7) can be simplified by using the auxiliary variables $\boldsymbol{\kappa}_c = \left[\kappa_{c,1}, \cdots, \kappa_{c,D}\right]$, and $\boldsymbol{\zeta}_p = \left[\zeta_{p,1}, \cdots, \zeta_{p,D}\right]$, respectively, as follows

$$\chi_{c,d} \geq \frac{\left|\overline{\mathbf{h}}_d^H \mathbf{w}_c\right|^2}{\kappa_{c,d}}, \quad \beta_{p,d} \geq \frac{\left|\overline{\mathbf{h}}_d^H \mathbf{w}_d\right|^2}{\zeta_{p,d}}, \quad \forall d \in \mathcal{D}, \tag{12}$$

where

$$\kappa_{c,d} \leq \sum\nolimits_{i=1}^{D} \left|\overline{\mathbf{h}}_d^H \mathbf{w}_i\right|^2 + \sigma_d^2, \tag{13a}$$

$$\zeta_{p,d} \leq \sum\nolimits_{i=1, i\neq d}^{D} \left|\overline{\mathbf{h}}_d^H \mathbf{w}_i\right|^2 + \sigma_d^2. \tag{13b}$$

The non-convex expressions in (12) can be further converted into an equivalent affine form utilizing the SCA technique as

$$\frac{2\Re\left(\mathbf{w}_c^{(n)H} \mathbf{h}_d \mathbf{h}_d^H \mathbf{w}_c\right)}{\kappa_{c,d}^{(n)}} - \frac{|\mathbf{h}_d^H \mathbf{w}_c^{(n)}|^2}{\kappa_{c,d}^{(n)2}} \geq \chi_{c,d}, \forall d \in \mathcal{D}, \tag{14a}$$

$$\frac{2\Re\left(\mathbf{w}_d^{(n)H} \mathbf{h}_d \mathbf{h}_d^H \mathbf{w}_d\right)}{\zeta_{p,d}^{(n)}} - \frac{|\mathbf{h}_d^H \mathbf{w}_d^{(n)}|^2}{\zeta_{p,d}^{(n)2}} \geq \beta_{p,d}, \forall d \in \mathcal{D}, \tag{14b}$$

where $\mathbf{w}_c^{(n)}$, $\mathbf{w}_d^{(n)}$, $\kappa_{c,d}^{(n)}$, and $\zeta_{p,d}^{(n)}$ are the values acquired at the n^{th} SCA iteration. Ultimately, design the transmit

beamforming in problem $\mathcal{P}1$ using n^{th} SCA iteration can be transformed as

$$\mathcal{P}2 \quad \max_{\substack{\mathbf{W},\boldsymbol{\varkappa}_c,\boldsymbol{\Upsilon}_p,\\ \boldsymbol{\beta}_p,\boldsymbol{\chi}_c,\boldsymbol{\kappa}_c,\boldsymbol{\zeta}_p}} \quad \sum\nolimits_{d=1}^{D}(\varkappa_{c,d}+\Upsilon_{p,d}) \tag{15a}$$

$$\text{s.t.} \quad (10e),(10g),(11b),(11c),(11d),(13a),(13b),(14a),(14b).$$

Due to its concave objective function and affine constraints, the simplified problem $\mathcal{P}2$ is converted into a standard convex problem [20], due to all optimization factors. Therefore, as shown in **Algorithm 1**, we provide an iterative approach for transmit beamforming design given the information on the convergence factor ϵ.

Algorithm 1 Design transmit Beamforming at BS using SCA

Input: $\overline{\psi}, \epsilon$, and $\mathbf{W}^0$

1: **Initialize** $n=0, R_{\text{Total}}^{(-1)}=0$
2: **while** $\left|R_{\text{Total}}^{(n)}-R_{\text{Total}}^{(n-1)}\right| \leq \epsilon$ or $n \leq N_{\max}^{\text{itr}}$ **do**,
3: Calculate $\kappa_{c,d}^{(n)}, \zeta_{p,d}^{(n)}, \chi_{c,d}^{(n)}, \beta_{p,d}^{(n)}$
4: $n=n+1$,
5: Solve $\mathcal{P}2$ to obtain the beamforming matrix $\mathbf{W}^{(n)}$,
6: Calculate $R_{\text{Total}}^{(n)}=\sum_{d=1}^{D}(\varkappa_{c,d}+\Upsilon_{p,d})$,
7: **end while**

Output: $\hat{R}_{\text{Total}}=R_{\text{Total}}^{(n)}$ and $\hat{\mathbf{W}}=\mathbf{W}^{(n)}$.

C. Passive RIS Beamforming Optimization

Similarly, the optimal passive RIS beamformer $\boldsymbol{\psi}$ is designed for the obtained $\mathbf{W}=\overline{\mathbf{W}}$. In such case, the optimization problem can be expressed by introducing the auxiliary variable $\tilde{\boldsymbol{\Upsilon}}_p, \tilde{\boldsymbol{\beta}}_p$, and $\tilde{\boldsymbol{\chi}}_c$. $\mathcal{P}0$ will be reformulated as

$$\mathcal{P}3 \quad \max_{\substack{\boldsymbol{\psi},\tilde{\boldsymbol{\varkappa}}_c,\\ \tilde{\boldsymbol{\Upsilon}}_p,\tilde{\boldsymbol{\beta}}_p,\tilde{\boldsymbol{\chi}}_c}} \quad \sum\nolimits_{d=1}^{D}(\tilde{\varkappa}_{c,d}+\tilde{\Upsilon}_{p,d}) \tag{16a}$$

$$\text{s.t.} \quad \sum\nolimits_{d=1}^{D}\tilde{\varkappa}_{c,d} \leq \log_2(1+\tilde{\chi}_{c,d}), \quad \forall d \in \mathcal{D}, \tag{16b}$$

$$\tilde{\varkappa}_{c,d}+\tilde{\Upsilon}_{p,d} \geq R_d^{min}, \quad \forall d \in \mathcal{D}, \tag{16c}$$

$$\tilde{\Upsilon}_{p,d} \leq \log_2(1+\tilde{\beta}_{p,d}), \quad \forall d \in \mathcal{D}, \tag{16d}$$

$$\tilde{\varkappa}_{c,d} \geq 0, \quad \forall d \in D, \tag{16e}$$

$$(6),(7),(10f),$$

where $\tilde{\boldsymbol{\Upsilon}}_p = \left[\tilde{\Upsilon}_{p,1},\cdots,\tilde{\Upsilon}_{p,D}\right]$, $\tilde{\boldsymbol{\beta}}_p = \left[\tilde{\beta}_{p,1},\cdots,\tilde{\beta}_{p,D}\right]$, $\tilde{\boldsymbol{\chi}}_c = \left[\tilde{\chi}_{c,1},\cdots,\tilde{\chi}_{c,D}\right]$. $\tilde{\boldsymbol{\varkappa}}_c = \left[\tilde{\varkappa}_{c,1},\cdots,\tilde{\varkappa}_{c,d}\right]$ is common rate allocation vector. The RIS beamformer design in $\mathcal{P}3$ is non-convex in nature. Therefore, we adopt the SCA method at fixed $\mathbf{W}=\overline{\mathbf{W}}$ and reformulate $\mathcal{P}3$ as

$$\mathcal{P}4 \quad \max_{\substack{\boldsymbol{\psi},\tilde{\boldsymbol{\varkappa}}_c,\tilde{\boldsymbol{\Upsilon}}_p,\\ \tilde{\boldsymbol{\beta}}_p,\tilde{\boldsymbol{\chi}}_c,\tilde{\boldsymbol{\kappa}}_c,\tilde{\boldsymbol{\zeta}}_p}} \quad \sum_{d=1}^{D}(\tilde{\varkappa}_{c,d}+\tilde{\Upsilon}_{p,d}) \tag{17a}$$

$$\text{s.t.} \quad \sum\nolimits_{i=1}^{D}\left|\mathbf{w}_i^H\mathbf{H}_A\text{diag}(\mathbf{f}_d(\mathbf{t}_d))\boldsymbol{\psi}\right|^2+\sigma_d^2 \leq \tilde{\kappa}_{c,d}, \tag{17b}$$

Algorithm 2 AO for Unified Solution.

Input: $\overline{\psi}, \overline{\mathbf{W}}, \mathcal{S}_{\max}^{\text{itr}}$ and **Initailize:** $\mathbf{W}^0, s=0$, and $R_{\text{Total}}^{-1}=0$

1: **while** $\left\|R_{\text{Total}}^{(s)}-R_{\text{Total}}^{(s-1)}\right\|^2 \leq \epsilon$ or $s \leq \mathcal{S}_{\max}^{\text{itr}}$ **do**,
2: Solve $\mathcal{P}4$ to obtain $\boldsymbol{\psi}^{(s+1)}$.
3: To obtain $\mathbf{W}^{(s+1)}$, solve $\mathcal{P}2$ using **Algorithm 1** with obtained $\boldsymbol{\psi}^{(s+1)}$ in step 2.
4: Update $s=s+1$.
5: **end while**

Output: $R_{\text{Total}}=R_{\text{Total}}^{(s)}$.

$$\sum\nolimits_{i=1,i\neq d}^{D}\left|\mathbf{w}_i^H\mathbf{H}_A\text{diag}(\mathbf{f}_d(\mathbf{t}_d))\boldsymbol{\psi}\right|^2+\sigma_d^2 \leq \tilde{\zeta}_{p,d} \tag{17c}$$

$$2\Re\Big(\Big(\overline{\mathbf{h}}_d^{\mathrm{R}}\boldsymbol{\psi}^{(q)}\Big)\overline{\mathbf{h}}_d^{\mathrm{R}}\boldsymbol{\psi}\Big)/\tilde{\kappa}_{c,d}^{(q)} - \left|\overline{\mathbf{h}}_d^{\mathrm{R}}\boldsymbol{\psi}^{(q)}\right|^2\tilde{\kappa}_{c,d}/\big(\tilde{\kappa}_{c,d}^{(q)}\big)^2 \geq \tilde{\chi}_{c,d}, \quad \forall d \in \mathcal{D}, \tag{17d}$$

$$2\Re\Big(\Big(\overline{\mathbf{h}}_d^{\mathrm{R}}\boldsymbol{\psi}^{(q)}\Big)\overline{\mathbf{h}}_d^{\mathrm{R}}\boldsymbol{\psi}\Big)/\tilde{\zeta}_{p,d}^{(q)} - \left|\overline{\mathbf{h}}_d^{\mathrm{R}}\boldsymbol{\psi}^{(q)}\right|^2\tilde{\zeta}_{p,d}/\big(\tilde{\zeta}_{p,d}^{(q)}\big)^2 \geq \tilde{\beta}_{p,d}, \quad \forall d \in \mathcal{D}, \tag{17e}$$

$$\sum\nolimits_{i=1}^{D}\log_2(1+\tilde{\chi}_{c,d}) \geq \tilde{\varkappa}_{c,d}, \quad \forall d \in \mathcal{D}, \tag{17f}$$

$$\log_2(1+\tilde{\beta}_{p,d}) \geq \tilde{\Upsilon}_{p,d}, \quad \forall d \in \mathcal{D}, \tag{17g}$$

$$|\psi_k| \leq 1, \quad k \in \mathcal{K},, \tag{17h}$$

$$(16b),(16c)(16d),(16e),$$

where $\tilde{\boldsymbol{\kappa}}_p = \left[\tilde{\kappa}_{p,1},\cdots,\tilde{\kappa}_{p,D}\right]$ and $\tilde{\boldsymbol{\zeta}}_p = \left[\tilde{\zeta}_{p,1},\cdots,\tilde{\zeta}_{p,D}\right]$ are the auxiliary variables and $\overline{\mathbf{h}}_d^{\mathrm{R}} = \mathbf{w}_i^H\mathbf{H}_A\text{diag}(\mathbf{f}_d(\mathbf{t}_d)), \forall d \in \mathcal{D}$. Here, $\boldsymbol{\psi}^{(q)}, \tilde{\kappa}_{c,d}^{(q)}, \tilde{\zeta}_{p,d}^{(q)}$ are the value of $\boldsymbol{\psi}, \tilde{\kappa}_{c,d}, \tilde{\zeta}_{p,d}$ obtained at q^{th} iteration, respectively. Note that the constraints in $\mathcal{P}4$ are convex. Thus, it is solved by adopting CVX [20].

D. Combined Approach and Computational Complexity

Finally, we propose an AO-based algorithm for joint design, as given in **Algorithm 2**. We use an iterative approach to alternately address the subproblem of transmit beamforming problem $\mathcal{P}2$ at BS and passive RIS beamformer design problem $\mathcal{P}4$. The output of each optimization problem in the current iteration is used as an input to the next iteration until a convergence ϵ is attained or up to iterations are concluded. Assume that the RIS phase-shift design converges within a maximum of $Q_{\max}^{\text{itr}}$ SCA iteration, and the precoder design solution in **Algorithm 1** convergence within $N_{\max}^{\text{itr}}$ iterations. The phase-shift design in $\mathcal{P}4$ has $8D+1$ constraints and $MD+6D$ variables. Therefore, $\mathcal{O}\Big(N_{\max}^{\text{itr}}(8D+1)^2(MD+6D)\Big)$ and $\mathcal{O}\Big(Q_{\max}^{\text{itr}}(8D+K)^2(K+6D)\Big)$ can be used to represent the algorithm's worst case complexity, respectively. Assuming that the suggested AO method converges with $\mathcal{S}_{\max}^{\text{itr}}$, the overall computational complexity in **Algorithm 2** is given by $\mathcal{O}\Big(S_{\max}^{\text{itr}}\Big(N_{\max}^{\text{itr}}(8D+1)^2(MD+6D)\Big)+\Big(Q_{\max}^{\text{itr}}(8D+K)^2(K+6D)\Big)\Big)$

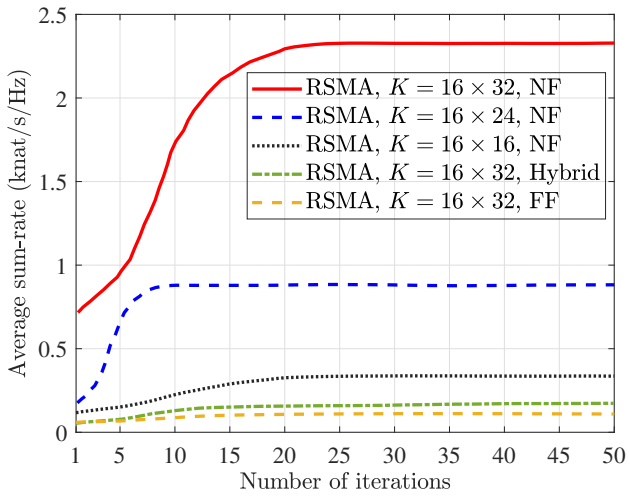

Fig. 2: Average sum-rate versus number of iterations.

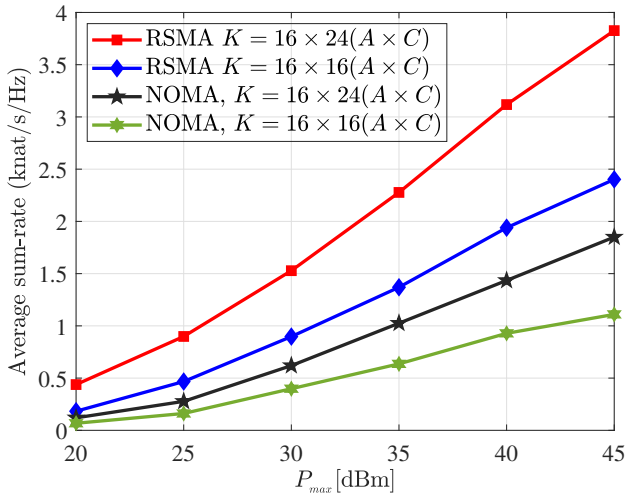

Fig. 3: Average sum-rate versus $P_{max}(D = 3, M = 20)$.

IV. Numerical Results

In this section, several numerical examples have been carried out to exhibit the efficacy and convergence of the proposed algorithm through Monte Carlo simulations, averaging over 10^5 independent realizations of randomly generated channels. The BS is located at (16.453m, 116.2630°,72.0471°) along with all users located at (4.1472m, 78.3022°,101.5349°), (8.6763m, 110.3402°,79.7452°), and (14.6328m, 101.1328°,119.5479°), respectively. The RIS is located at the center of the coordinate (0, 0°, 0°). Further, unless otherwise specified, $\sigma_d^2 = -90$dBm, $D = 3$, and $M = 20$. Furthermore, the operating frequency is considered to 24GHz with a HARIS aperture length of 1m. Therefore, the Fresnel and Rayleigh distance works out to be 1.3m and 160.2m, respectively. Therefore, any radio receiver lying within this range will experience the effect of near-field [18], [19], [21]. Hence, all the users considered lie within the near-field region of the RIS-aided holographic MIMO.

Fig. 2 illustrates the convergence characteristics of **Algorithm2**, which is obtained with P_{max} = 25 dBm, M = 20, and

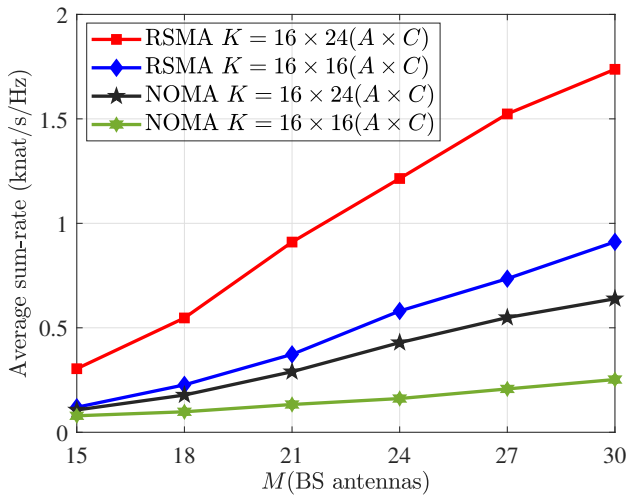

Fig. 4: Average sum-rate versus $M(D = 3, P_{max} = 15$dBm).

RIS elements $K(A \times C) = 256, 384, 512$. It is evident that the algorithm converges after a few iterations, demonstrating its effectiveness. Furthermore, the average sum-rate improves as K increases. This is because the transmitted signal is transformed into the enhanced beam over the spatially continuous aperture of the holographic MIMO. Moreover, we compared the considered near-field assumptions with the far-field and hybrid[1] channel assumptions. We observed that, compared to the far-field and hybrid systems, the near-field assumptions are more suitable for such scenarios due to the fact that all the nodes lies within the Rayleigh distance w.r.t. RIS.

Fig. 3 depicts the average sum-rate (in knat/s/Hz) versus the total transmit power P_{max} for different numbers of RIS elements (K). It can be observed that the performance increases with increasing P_{max}. The main reason for this is the increase in net SINR corresponding to BS at each user due to the rise in the maximum available power for transmission at BS. Furthermore, one can also observe that the use of RSMA can significantly enhance the performance compared to the NOMA. This is because the RSMA provides the ability to split the rate in to common and private parts at the transmitter and combining the same at the receiver. This splitting in RSMA, when combined with optimum resource allocation, leads to a higher achievable rate at each user compared to NOMA, especially under high transmit power.

Fig. 4 shows the average sum-rate versus the number of the BS antennas for different numbers of RIS elements. It can be observed that the performance increases with the increase in the number of BS antennas. The reason for this is that, by increasing the number of antennas, the channel diversity gain increases, thus improving the average sum-rate. Further, it can also be seen that the RSMA scheme outperform the NOMA.

Fig. 5 illustrates the average sum-rate versus the number of users (D) for different number of RIS elements with $M = 20, P_{max} = 30$dBm. It can be observed that the average

[1]The hybrid-field assumptions are adopted according to the considered system model following [22].

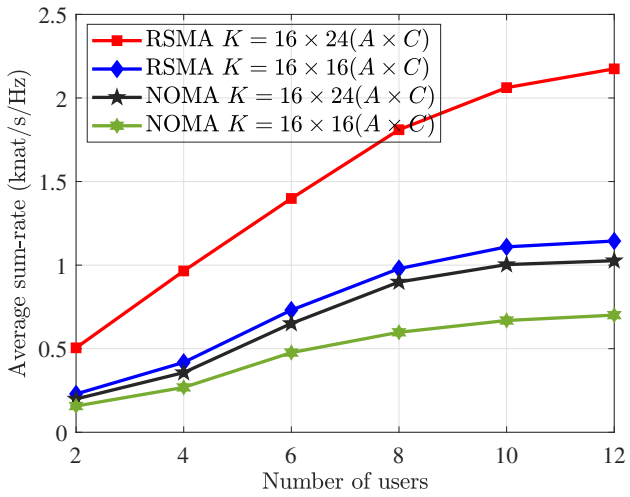

Fig. 5: Average sum-rate versus D.

sum-rate increases with increase in D. It can slo be seen that, the RSMA system achieves better performance as compared to the NOMA scheme. This is because the RSMA provides the ability to split the rate in to common and private parts at the transmitter and combining the same at the receiver. This splitting in RSMA, when combined with optimum resource allocation, leads to a higher achievable rate at each user compared to NOMA, especially under high transmit power. Hence, with increasing D, the difference between average sum rate of RSMA and NOMA also increases significantly.

V. Conclusion

In this work, we proposed the use of RSMA in a RIS-aided holographic MIMO-based wireless network under the near-field framework. We discussed an analytical framework wherein an AO-based algorithm hat effectively solves the joint optimization problem using analytical tools like successive convex approximation technique. Besides, several numerical simulations were performed to validate the efficacy and superiority of the proposed algorithm. The fast convergence of the algorithm, despite the variation in the number of elements, ensures the efficacy of the proposed algorithm. Furthermore, we also observed that increasing the number of passive elements or power at RIS results in better system performance. Moreover, it was shown that, for the given set of parameters, the RSMA-based system outperformed the NOMA system.

References

[1] Z. Liwen, F. Qamar, M. Liaqat, M. Nour Hindia, and K. Akram Zainol Ariffin, "Towards Efficient 6G IoT Networks: A Perspective on Resource Optimization Strategies, Challenges, and Future Directions," *IEEE Access*, May 2024.

[2] P. Yang, Y. Xiao, M. Xiao, and S. Li, "6G wireless communications: Vision and potential techniques," *IEEE Netw.*, vol. 33, no. 4, pp. 70–75, 2019.

[3] C. Huang, S. Hu, G. C. Alexandropoulos, A. Zappone, C. Yuen, R. Zhang, M. D. Renzo, and M. Debbah, "Holographic MIMO surfaces for 6G wireless networks: Opportunities, challenges, and trends," *IEEE Wireless Commun.*, vol. 27, no. 5, pp. 118–125, 2020.

[4] A. Pizzo, T. L. Marzetta, and L. Sanguinetti, "Spatially-stationary model for holographic MIMO small-scale fading," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 9, pp. 1964–1979, 2020.

[5] S. Zeng, H. Zhang, B. Di, H. Qin, X. Su, and L. Song, "Reconfigurable refractive surfaces: An energy-efficient way to holographic MIMO," *IEEE Commun. Lett.*, vol. 26, no. 10, pp. 2490–2494, 2022.

[6] C. You, Y. Cai, Y. Liu, M. D. Renzo, T. M. Duman, A. Yener, and A. L. Swindlehurst, "Next generation advanced transceiver technologies for 6G," 2024. [Online]. Available: https://arxiv.org/abs/2403.16458

[7] S. Kurma, T. A. Lestari, K. Singh, A. Paul, and S. Mumtaz, "Active RIS in digital twin-based URLLC IoT networks: Fully-connected versus sub-connected?" *IEEE Trans. Wireless Commun.*, vol. 23, no. 9, pp. 12 354–12 367, Sep. 2024.

[8] J. Xu, L. You, G. C. Alexandropoulos, X. Yi, W. Wang, and X. Gao, "Near-field wideband extremely large-scale MIMO transmissions with holographic metasurface-based antenna arrays," *IEEE Trans. Wireless Commun.*, 2024.

[9] N. J. Myers and R. W. Heath, "Infocus: A spatial coding technique to mitigate misfocus in near-field LoS beamforming," *IEEE Trans. Wireless Commun.*, vol. 21, no. 4, pp. 2193–2209, Apr. 2022.

[10] H. Zhang, N. Shlezinger, F. Guidi, D. Dardari, and Y. C. Eldar, "6G wireless communications: From far-field beam steering to near-field beam focusing," *IEEE Commun. Mag.*, vol. 61, no. 4, pp. 72–77, Mar. 2023.

[11] W. Chen, Z. Wei, Z. Yang, D. W. Kwan Ng, and M. Matthaiou, "Beamforming design for RIS-aided MIMO communication: A piece-wise near-field model," in *Proc. IEEE ICCC Workshops*, Aug. 2024, pp. 385–390.

[12] S. Lv, Y. Liu, X. Xu, A. Nallanathan, and A. L. Swindlehurst, "RIS-aided near-field MIMO communications: Codebook and beam training design," *IEEE Trans. Wireless Commun.*, vol. 23, no. 9, pp. 12 531–12 546, Sep. 2024.

[13] Y. Liu, X. Mu, X. Liu, M. Di Renzo, Z. Ding, and R. Schober, "Reconfigurable intelligent surface-aided multi-user networks: Interplay between NOMA and RIS," *IEEE Wireless Commun.*, vol. 29, no. 2, pp. 169–176, 2022.

[14] 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 J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.

[15] M. Katwe, K. Singh, B. Clerckx, and C.-P. Li, "Rate splitting multiple access for energy efficient RIS-aided multi-user short-packet communications," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2022, pp. 644–649.

[16] X. Li, T. Wang, H. Tong, Z. Yang, Y. Mao, and C. Yin, "Sum-rate maximization for active RIS-aided downlink RSMA system," in *Proc. IEEE INFOCOM WKSHPS*, May 2023, pp. 1–6.

[17] G. Zheng, M. Wen, J. Wen, and C. Shan, "Joint hybrid precoding and rate allocation for RSMA in near-field and far-field massive MIMO communications," *IEEE Wireless Commun. Lett.*, vol. 13, no. 4, pp. 1034–1038, Apr. 2024.

[18] Y. Liu, Z. Wang, J. Xu, C. Ouyang, X. Mu, and R. Schober, "Near-field communications: A tutorial review," *IEEE Open. J. Commun. Soc.*, vol. 4, pp. 1999–2049, 2023.

[19] N. Song and H. Dai, "Optimal beam-focusing design for 6g near-field swipt systems," *IEICE Commun. Express*, vol. 13, no. 3, pp. 84–87, 2024.

[20] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[21] Z. Wang, X. Mu, and Y. Liu, "Near-field integrated sensing and communications," *IEEE Commun. Lett.*, vol. 27, no. 8, pp. 2048–2052, 2023.

[22] X. Wei and L. Dai, "Channel estimation for extremely large-scale massive MIMO: Far-field, near-field, or hybrid-field?" *IEEE Commun. Lett.*, vol. 26, no. 1, pp. 177–181, Jan. 2022.