

Robust Beamforming Design for STAR-RIS Aided RSMA Network with Hardware Impairments

Ziyue Wang, Xiaoyan Ma, Xingyu Peng, Zheao Li, Jinyuan Liu, Yongliang Guan, *Senior Member, IEEE*,
and Chau Yuen, *Fellow, IEEE*

Abstract—In this article, we investigate the robust beamforming design for a simultaneous transmitting and reflecting reconfigurable intelligent surface (STAR-RIS) aided downlink rate-splitting multiple access (RSMA) communication system, where both transceivers and STAR-RIS suffer from the impact of hardware impairments (HWI). A base station (BS) is deployed to transmit messages concurrently to multiple users, utilizing a STAR-RIS to improve communication quality and expand user coverage. We aim to maximize the achievable sum rate of the users while ensuring the constraints of transmit power, STAR-RIS coefficients, and the actual rate of the common stream for all users. To solve this challenging high-coupling and non-convexity problem, we adopt a fractional programming (FP)-based alternating optimization (AO) approach, where each sub-problem is addressed via successive convex approximation (SCA) and penalty function (PF) methods. Numerical results demonstrate that the proposed scheme outperforms other multiple access schemes and conventional passive RIS in terms of the achievable sum rate. Additionally, considering the HWI of the transceiver and STAR-RIS makes our algorithm more robust than when such considerations are not included.

Index Terms—Rate-splitting multiple access (RSMA), simultaneous transmitting and reflecting reconfigurable intelligent surface (STAR-RIS), hardware impairments (HWI), phase noise, fractional programming (FP).

I. INTRODUCTION

THE rapid expansion of the Internet of Things (IoT) has driven an exponential growth in the number of mobile users and network traffic volumes over the past few decades [1], [2]. Given the limited availability of spectrum resources, next-generation communication networks are urgently seeking new technologies to enhance spectrum efficiency, aiming to achieve higher throughput, enhanced reliability, expanded support for machine-type communications, and broader coverage for both communication and sensing. As a result, researchers

are increasingly focused on developing innovative solutions such as network slicing and the non-orthogonal multiple access (NOMA) [3], [4]. These advancements are designed to maximize resource utilization, provide seamless connectivity, and enable scalable, flexible networks capable of supporting diverse IoT applications.

Traditional multiple access techniques, such as orthogonal multiple access (OMA), provide feasible solutions for resource distribution in existing multi-user communication systems. However, OMA faces limitations in terms of low spectral efficiency and high-complexity orthogonality restoring measurements [5]. Unlike traditional OMA schemes, which allocate resources based on time or frequency, NOMA differentiates users by varying their power levels within the same time and frequency resource block [6]–[8]. In NOMA, each message of the user is encoded into distinct streams, which are superimposed and transmitted simultaneously to all users. At the receiver side, successive interference cancellation (SIC) is used to decode the individual messages based on a predefined decoding order. Previous studies have demonstrated that NOMA provides improved spectral efficiency and enhanced user fairness in a variety of communication scenarios [9], [10]. However, NOMA does not fully exploit the spatial dimension. In [11], it has been shown that in multiple-input multiple output (MIMO) scenarios, the sum multiplexing gain of space division multiple access (SDMA) is always no less than NOMA. This suggests that NOMA underperforms in terms of sum rate compared to SDMA at high signal-to-noise ratio (SNR) levels, indicating its limited ability to fully exploit the multi-antenna capabilities of communication systems.

Recently, rate-splitting multiple access (RSMA) has emerged as a promising technology for next-generation networks [12]. Unlike NOMA and traditional SDMA, the key concept of RSMA is to split the message of each user into a common part and a private part. The common parts of all users are combined and encoded into a common message, while the private parts are separately encoded into individual messages. At the receiver side, users first decode the common message while treating all private streams as interference. Once the common message is successfully decoded, each user then proceeds to decode its own private message, treating the remaining signals from other users as interference. This approach allows interference to be partially treated as noise. By dynamically adjusting the common stream and transmission power, RSMA can adapt to various interference scenarios

This work was supported in part by the A*STAR (Agency for Science, Technology and Research) Singapore, under Grant No. M22L1b0110; and in part by the MOE (Ministry of Education, Singapore), under MOE Tier 2 Award number T2EP50124-0032. (*Corresponding author: Chau Yuen.*)

Ziyue Wang, Zheao Li, Jinyuan Liu, Yongliang Guan and Chau Yuen are with School of Electrical and Electronics Engineering, Nanyang Technological University, 639798, Singapore (Email: wang1953@e.ntu.edu.sg; zheao001@e.ntu.edu.sg; jinyuan001@e.ntu.edu.sg; EYLGuan@ntu.edu.sg; Chau.Yuen@ntu.edu.sg).

Xiaoyan Ma is with School of Electrical and Computer Engineering, Purdue University, USA (Email: ma946@purdue.edu).

Xingyu Peng is with College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China, and Zhejiang Provincial Key Laboratory of Info. Proc., Commun. & Netw. (IPCAN), Hangzhou 310027, China (Email: peng_xingyu@zju.edu.cn).

and further improve spectrum efficiency compared to existing NOMA or SDMA schemes [13]-[15].

Although RSMA has an outstanding and flexible capability for resource allocation, which can enhance spectrum efficiency and resource utilization to a certain extent, the high frequency characteristic in the future network will lead to severe signal fading and poor traversal capability despite the denser and more diverse base station (BS) deployments, which will limit the full potential of RSMA. To further capitalize on the advantages of RSMA in next-generation networks, the emergence of reconfigurable intelligent surface (RIS) can help to overcome this shortcoming [16]. RIS is a reconfigurable planar array composed of numerous passive elements. By controlling the reflection unit array on the RIS surface via a microcontroller, RIS can reflect incoming signals in specific directions, actively shaping the channel environment, enhancing line-of-sight (LoS) components and mitigating signal blind spots. The passive nature of RIS that requires no additional energy beyond the micro-controller, makes it not only cost-effective and easy to deploy, but also low-maintenance. Such characteristics endow RIS with broad application prospects in multiple aspects of communication systems, making it a highly promising technology for next-generation networks [17]-[25]. Specifically, in the RSMA transmission scheme, deploying RIS can effectively enhance system performance by providing additional direct transmission paths, which improves both the received signal power at users and the degrees of freedom for beamforming design. Previous research has demonstrated the value of RIS aided RSMA systems. The work [26] addressed the maximization of the minimum group rate in multigroup multicast communication scenarios. The work [27] focused on maximizing the weighted sum rate for multi-user systems, while the work [28] proposed a robust beamforming design for downlink RIS aided symbiotic radio systems with RSMA. The above studies show superior performance for RIS aided RSMA scenarios compared to those without RIS aided.

Traditional RIS can only enable half-space coverage. In fact, current RIS prototypes are limited by hardware and can only serve angles of less than 180 degrees [29]. Simultaneous transmitting and reflecting reconfigurable intelligent surfaces (STAR-RIS), a new type of RIS theoretically capable of serving full-space coverage, offers mode switching, energy splitting, and time switching modes [30]. The hardware implementation of STAR-RIS is different with the conventional passive RIS, as well as its signal model [31] [32]. In the energy-splitting (ES) mode, each element on the STAR-RIS surface has four configurable degrees of freedom to adjust the incident signal: the amplitude and phase of both the reflection and transmission regions. This represents a significant increase in system degrees of freedom and versatility compared to conventional RIS [33]. It surpasses traditional passive RIS in both spatial coverage freedom and the feasible domain of optimization algorithms.

The integrated STAR-RIS and RSMA system, with its 360-degree flexible signal enhancement and excellent interference management, is poised to become a new architectural paradigm for next-generation communication networks. In [34], the authors demonstrated that STAR-RIS empowered

RSMA systems achieve lower BS energy consumption at the same outage probability compared to NOMA systems. In [35], the authors proposed a proximal policy optimization (PPO)-based deep reinforcement learning (DRL) algorithm to solve challenging non-convex optimization problems. The results show that compared to conventional RIS, STAR-RIS significantly enhances system sum rate and user fairness, and the integrated STAR-RIS and RSMA system achieves more notable performance gains than deploying NOMA methods. In [36], the authors considered the physical layer security performance of downlink integrated STAR-RIS and RSMA systems, using a successive convex approximation (SCA)-based penalization algorithm to achieve significantly enhanced secrecy rates compared to conventional RIS. In [37], it was demonstrated that STAR-RIS aided RSMA systems also perform excellently in integrated communication and sensing.

Despite previous studies highlighting the potential of STAR-RIS aided RSMA systems, perfect hardware implementations at both the transceiver and the STAR-RIS side are typically assumed. In practical communication system applications, hardware degrades over time due to environmental influences, leading to signal distortions that deviate from the intended targets. This degradation affects key components such as power amplifiers, digital-to-analog converters, oscillators, as well as the STAR-RIS itself [38]. Specifically, for STAR-RIS, the hardware impairments (HWI) are usually seen as the phase noise among the elements [39], which is generated either by phase estimation error or by quantization error. HWI can degrade system performance [40]-[42] and affect the effective multi-user management capabilities of RSMA. Therefore, to ensure robust system performance under realistic deployment conditions, it is necessary to consider the impact of HWI during signal modeling and compensate for their effects during optimization processes using signal processing algorithms. However, considering HWI at the transceiver and STAR-RIS introduces additional variables in the optimization process, complicating signal derivation and increasing algorithmic overhead, making research on HWI consideration more challenging.

In summary, HWI introduces deviations in the signal processing procedure under practical conditions, negatively affecting algorithm performance and significantly degrading system efficiency. Although integrated STAR-RIS and RSMA system offers substantial advantages in various aspects, existing research has not yet fully addressed the impact of HWI in such systems. To bridge this gap, this paper investigates a downlink multi-user scenario that comprehensively incorporates HWI, including impairments at the transceivers and phase noise at the STAR-RIS. This modeling approach better aligns with practical engineering applications and enables optimization and compensation for HWI within the proposed algorithm. The main contributions of this paper are summarized as follows:

- We investigate a STAR-RIS aided downlink RSMA communication system considering HWI at both the transceivers and the STAR-RIS side. In this setup, the BS transmits both private and common messages to multiple users using RSMA with the help of the STAR-RIS. Then, we aim to maximize the achievable sum rate for the

proposed STAR-RIS aided RSMA robust beamforming design by jointly optimizing the transmit beamforming vector at the BS side, the STAR-RIS coefficients, and the actual transmission rate for the common stream. By accounting the HWI, we enhance the STAR-RSMA system to better reflect real-world conditions, compensating for the progressive signal distortion caused by environmental factors over time.

- To tackle the high coupling and non-convexity optimization problem, we first employ fractional programming (FP) to reformulate the original objective function into a more tractable form. The resulting problem is then decomposed into several sub-problems within an alternating optimization (AO) framework. Furthermore, the transmit beamforming vector and the STAR-RIS coefficients are jointly optimized using a combination of SCA and penalty function (PF) algorithms.
- Simulation results demonstrate that the proposed FP-AO method achieves rapid convergence and delivers superior performance in the considered scenario. Furthermore, it is verified that the deployment of STAR-RIS significantly outperforms conventional RIS in terms of achievable sum rate performance, yielding up to 23% higher gain in optimal situations. The robust beamforming design maintains consistent performance advantages over non-robust designs under various HWI conditions. Additionally, RSMA scheme leads to a notable enhancement in communication quality as compared to the NOMA and OMA scheme, while demonstrating stronger resilience against HWIs induced by both transceiver chains and reflecting elements.

Notation: In this paper, the notations used are shown as follows. Scalars are denoted by lowercase, vectors in bold lower case, and matrices in bold upper case letters. For the vector $\mathbf{x}$, x_n , $|\mathbf{x}|$ and $\text{diag}(\mathbf{x})$ represent its n -th element, the absolute value of the vector, and the diagonal matrix formed by the vector, respectively. For the matrix $\mathbf{X}$, $\mathbf{X}^T$, $\mathbf{X}^H$, $\text{Tr}(\mathbf{X})$, $|\mathbf{X}|_{n,n}$ denotes its transpose, conjugate transpose, trace, and the element at the (n, n) -th position of the matrix. The expectation operator is represented by $\mathbb{E}[\cdot]$. $\odot$ denotes the Hadamard product. $\mathbb{C}^{a \times b}$ represents the set of all complex matrices of size $a \times b$. diag represents a diagonal matrix whose diagonal entries are the same as those of the input matrix.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

Fig. 1 shows a STAR-RIS aided downlink RSMA system where the BS is equipped with M antennas, and a STAR-RIS consisting of N elements provides additional signal paths for the users. The STAR-RIS is connected to the BS via a micro-controller, with each element operating in the ES mode. Specifically, each unit on STAR-RIS is capable of simultaneously reflecting and transmitting signals while the power of the output signal is equal to the incident signal, theoretically. Assume there are K users distributed across both the reflection and transmission region of STAR-RIS. Let $\mathcal{T} = \{1, 2, \dots, k_0\}$ and $\mathcal{R} = \{k_0 + 1, k_0 + 2, \dots, K\}$ represent the

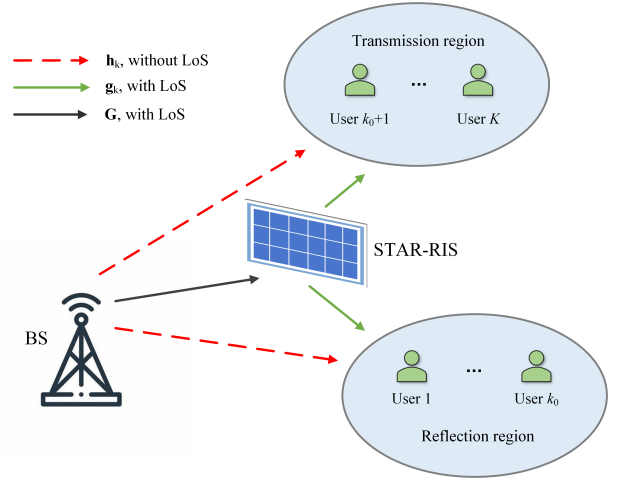


Fig. 1. A STAR-RIS aided RSMA communication system with HWI.

users in reflection region and transmission region, respectively, where $k_0 < K$ make sure both region have user located. Then let $\mathcal{K} = \mathcal{T} \cup \mathcal{R}$ be the set of the whole users. By adopting RSMA scheme, for any $k \in \mathcal{K}$, the message sent to the k -th user is divided into a common part m_k^c and a private part m_k^p . The transmitter uses the codebook shared by the BS and all users to merge the common parts m_k^c , into a single stream s_c to ensure all the users can decode the message from the common stream. For each m_k^p , a codebook known only to the BS and the k -th user is used to encode it into a stream s_k . Therefore, the BS will continuously transmit a common stream s_c and K private streams s_k at the same time. To assess the theoretical security capacity of the system, we consider a scenario where complete CSI is available. Given that this assumption can be flexibly relaxed in practical applications, the findings obtained in this paper serve as an upper-bound for practical scenarios. The relevant channel estimation techniques for STAR-RIS aided systems can be found in [43]. Therefore, the transmit signal can be expressed as

$$\mathbf{x} = \sum_{k=1}^K \mathbf{f}_k s_k + \mathbf{f}_c s_c + \mathbf{m}_t, \quad (1)$$

where $\mathbf{f}_k \in \mathbb{C}^{M \times 1}$, $\mathbf{f}_c \in \mathbb{C}^{M \times 1}$ represents the beamforming vector of private stream and common stream. $\mathbf{m}_t$ is the HWI variable produced by the transmitter. Previous research have done sufficient investigation both theoretically and practically on this variable considering the joint effect of the component in RF chain and show that it can be modeled as $\mathbf{m}_t \sim \mathcal{CN}(\mathbf{0}, \mu_t \text{diag}(\sum_{k=1}^K \mathbf{f}_k \mathbf{f}_k^H + \mathbf{f}_c \mathbf{f}_c^H))$ [42], where $\mu_t \geq 0$ is the ratio of transmit distorted noise power to transmit signal power. Let $\mathbb{E}[s_k s_k^H] = 1, \forall k \in \mathcal{K}$.

The channel from BS to the k -th user, BS to STAR-RIS and STAR-RIS to the k -th user are denoted by $\mathbf{h}_k \in \mathbb{C}^{M \times 1}$, $\mathbf{G} \in \mathbb{C}^{N \times M}$ and $\mathbf{g}_k \in \mathbb{C}^{N \times 1}$, where the $\mathbf{h}_k$ is without LoS component.¹ The signal received by the k -th user can be

¹The algorithm proposed in this work can be extended to general channel conditions, including scenarios where a LoS component exists between the BS and users.

$$\text{SINR}_{p,k} = \frac{|\mathbf{t}_k^H \mathbf{f}_k|^2}{\sum_{j \neq k} |\mathbf{t}_k^H \mathbf{f}_j|^2 + \mu_r (\sum_{j=1}^K |\mathbf{t}_k^H \mathbf{f}_j|^2 + |\mathbf{t}_k^H \mathbf{f}_c|^2) + (1 + \mu_r) \mu_t \mathbf{t}_k^H \text{diag}(\sum_{j=1}^K \mathbf{f}_j \mathbf{f}_j^H + \mathbf{f}_c \mathbf{f}_c^H) \mathbf{t}_k + (1 + \mu_r) \sigma_k^2}, \quad (7)$$

$$\text{SINR}_{c,k} = \frac{|\mathbf{t}_k^H \mathbf{f}_c|^2}{\sum_{j=1}^K |\mathbf{t}_k^H \mathbf{f}_j|^2 + \mu_r (\sum_{j=1}^K |\mathbf{t}_k^H \mathbf{f}_j|^2 + |\mathbf{t}_k^H \mathbf{f}_c|^2) + (1 + \mu_r) \mu_t \mathbf{t}_k^H \text{diag}(\sum_{j=1}^K \mathbf{f}_j \mathbf{f}_j^H + \mathbf{f}_c \mathbf{f}_c^H) \mathbf{t}_k + (1 + \mu_r) \sigma_k^2}, \quad (8)$$

represented as

$$y_k = \tilde{y}_k + m_{r,k}, \quad (2)$$

where $m_{r,k} \sim \mathcal{CN}(0, \mu_{r,k} \mathbb{E}\{|\tilde{y}_k|^2\})$ denotes the HWI at the k -th user. $\mu_{r,k} \geq 0$ is the ratio of distorted noise power to received signal power at the k -th user and $\tilde{y}_k$ represents the corresponding received signal:

$$\tilde{y}_k = \begin{cases} (\mathbf{h}_k^H + \mathbf{g}_k^H \Phi_t \mathbf{E}_t \mathbf{G}) \mathbf{x} + n_k, & k \in \mathcal{T}, \\ (\mathbf{h}_k^H + \mathbf{g}_k^H \Phi_r \mathbf{E}_r \mathbf{G}) \mathbf{x} + n_k, & k \in \mathcal{R}, \end{cases} \quad (3)$$

$\Phi_i = \text{diag}(\mathbf{v}_i)$ represents the diagonal reflection matrix of the STAR-RIS, where $\mathbf{v}_i = [\sqrt{\beta_1^i} e^{j\theta_1^i}, \sqrt{\beta_2^i} e^{j\theta_2^i}, \dots, \sqrt{\beta_N^i} e^{j\theta_N^i}]^T$ with $i = \{t, r\}$ represents the transmission region and reflection region, $\sqrt{\beta_n^i}$ and θ_n being the amplitude and the phase of the n -th element, respectively. $\mathbf{E}_i = \text{diag}(\phi_i)$ with $\phi_i = [e^{j\tilde{\theta}_1^i}, e^{j\tilde{\theta}_2^i}, \dots, e^{j\tilde{\theta}_N^i}]^H$, $i = \{t, r\}$ being the phase error vector. It is assumed $\tilde{\theta}_n^i, n = 1, 2, \dots, N$ is uniformly distributed in $[-\frac{\pi}{2}, \frac{\pi}{2}]$ [40], [44]. $n_k \sim \mathcal{CN}(0, \sigma_k^2)$ is the additive white gaussian noise (AWGN).

By applying RSMA scheme, we first denote

$$\mathbf{t}_k^H = \begin{cases} \mathbf{h}_k^H + \mathbf{g}_k^H \Phi_t \mathbf{E}_t \mathbf{G}, & k \in \mathcal{T}, \\ \mathbf{h}_k^H + \mathbf{g}_k^H \Phi_r \mathbf{E}_r \mathbf{G}, & k \in \mathcal{R}, \end{cases} \quad (4)$$

as the total channel of the k -th user. Hence, the achievable rate for the common stream of user k can be expressed as

$$R_{p,k} = \log_2(1 + \text{SINR}_{p,k}). \quad (5)$$

For private stream, the achievable rate of user k can be expressed as

$$R_{c,k} = \log_2(1 + \text{SINR}_{c,k}), \quad (6)$$

where the $\text{SINR}_{c,k}$ and $\text{SINR}_{p,k}$ are given in (7) and (8) at the top of the page.

Due to the randomness of the phase matrix generated by the HWI of STAR-RIS, the actual data rate is hard to obtain directly. To tackle this problem, we use the average rate here to approximate $R_{p,k}$ and $R_{c,k}$ [38], [45], [46] by taking expectation to the phase noise matrix $\mathbf{E}_i$ in (7) and (8). The average achievable rate can be represented as

$$R_{j,k} \approx \tilde{R}_{j,k} = \log_2(1 + \tilde{\gamma}_{j,k}), j = \{p, c\}. \quad (9)$$

Since $\mathbf{t}_k$ is the only variable which contains $\mathbf{E}_i$, we first address $\mathbb{E}\{\mathbf{t}_k \mathbf{t}_k^H\}$ as follows:

$$\begin{aligned} \mathbb{E}\{\mathbf{t}_k \mathbf{t}_k^H\} &= \mathbf{G}^H \Phi_t^H \text{diag}(\mathbf{g}_k) \mathbb{E}\{\phi_t \phi_t^H\} \text{diag}(\mathbf{g}_k^H) \Phi_t \mathbf{G} \\ &+ \mathbf{G}^H \Phi_t^H \text{diag}(\mathbf{g}_k) \mathbb{E}\{\phi_t^*\} \mathbf{h}_k^H + \mathbf{h}_k \mathbb{E}\{\phi_t^*\} \text{diag}(\mathbf{g}_k^H) \Phi_t \mathbf{G} \\ &+ \mathbf{h}_k \mathbf{h}_k^H, k \in \mathcal{T}, \end{aligned} \quad (10)$$

$$\begin{aligned} \mathbb{E}\{\mathbf{t}_k \mathbf{t}_k^H\} &= \mathbf{G}^H \Phi_r^H \text{diag}(\mathbf{g}_k) \mathbb{E}\{\phi_r \phi_r^H\} \text{diag}(\mathbf{g}_k^H) \Phi_r \mathbf{G} \\ &+ \mathbf{G}^H \Phi_r^H \text{diag}(\mathbf{g}_k) \mathbb{E}\{\phi_r^*\} \mathbf{h}_k^H + \mathbf{h}_k \mathbb{E}\{\phi_r^*\} \text{diag}(\mathbf{g}_k^H) \Phi_r \mathbf{G} \\ &+ \mathbf{h}_k \mathbf{h}_k^H, k \in \mathcal{R}. \end{aligned} \quad (11)$$

In order to calculate $\mathbb{E}\{\phi_i \phi_i^H\}$ and $\mathbb{E}\{\phi_i^*\}$, we denote $\delta_\theta = \tilde{\theta}_i - \tilde{\theta}_j, \forall i, j = 1, 2, \dots, N$. The probability function of $\tilde{\theta}_n$ can be express as $f(\tilde{\theta}_n) = \frac{1}{\pi}$. Thus,

$$\delta_\theta = \begin{cases} \frac{1}{\pi^2} \delta_\theta + \frac{1}{\pi}, & \delta_\theta \in [-\pi, 0], \\ -\frac{1}{\pi^2} \delta_\theta + \frac{1}{\pi}, & \delta_\theta \in [0, \pi], \end{cases} \quad (12)$$

and since we have $\mathbb{E}_{\delta_\theta}\{e^{j\delta_\theta}\} = \mathbb{E}_{\delta_\theta}\{e^{j\theta_i - j\theta_j}\} = \int_{-\pi}^{\pi} f(\delta_\theta) e^{j\delta_\theta} d\delta_\theta = \frac{4}{\pi^2}$, we can give the expression

$$\mathbb{E}\{\phi_i \phi_i^H\} = \mathbf{I}_N + \mathbf{J} = \begin{pmatrix} \frac{1}{\pi^2} & \dots & \frac{4}{\pi^2} \\ \frac{4}{\pi^2} & \dots & \frac{4}{\pi^2} \\ \vdots & \ddots & \vdots \\ \frac{4}{\pi^2} & \dots & 1 \end{pmatrix}, \quad (13)$$

where $\mathbf{J}$ is a $N \times N$ matrix

$$\mathbf{J}_{(i_1, i_2)} = \begin{cases} 0, & i_1 = i_2, \\ \frac{4}{\pi^2}, & i_1 \neq i_2. \end{cases} \quad (14)$$

For $\mathbb{E}\{\phi_i^*\}$, we have $\mathbb{E}_{\theta_i}\{e^{-j\theta_i}\} = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\theta_i) (\cos \theta_i - j \sin \theta_i) d\theta_i = \frac{2}{\pi}$. Then, we can have

$$\mathbb{E}\{\phi_i^*\} = \frac{2}{\pi} \mathbf{1}_N, \quad (15)$$

where $\mathbf{1}_N$ is a column vector with all elements being one. By applying (13) and (15), (10) and (11) can be replaced by

$$\begin{aligned} \mathbb{E}\{\mathbf{t}_k \mathbf{t}_k^H\} &= \begin{cases} \mathbf{G}^H \Phi_t^H \text{diag}(\mathbf{g}_k) \mathbf{D} \mathbf{D}^T \text{diag}(\mathbf{g}_k^H) \Phi_t \mathbf{G} + \\ (\frac{2}{\pi} \mathbf{G}^H \Phi_t^H \mathbf{g}_k + \mathbf{h}_k) (\frac{2}{\pi} \mathbf{g}_k^H \Phi_t \mathbf{G} + \mathbf{h}_k^H), k \in \mathcal{T} \\ \mathbf{G}^H \Phi_r^H \text{diag}(\mathbf{g}_k) \mathbf{D} \mathbf{D}^T \text{diag}(\mathbf{g}_k^H) \Phi_r \mathbf{G} + \\ (\frac{2}{\pi} \mathbf{G}^H \Phi_r^H \mathbf{g}_k + \mathbf{h}_k) (\frac{2}{\pi} \mathbf{g}_k^H \Phi_r \mathbf{G} + \mathbf{h}_k^H), k \in \mathcal{R} \end{cases} \\ &= \bar{\mathbf{T}}_k, \end{aligned} \quad (16)$$

where $\mathbf{D} \mathbf{D}^T \triangleq (1 - \frac{4}{\pi^2}) \mathbf{I}_N$. Now, the exact expressions of $\tilde{\gamma}_{p,k}$ and $\tilde{\gamma}_{c,k}$ can be expressed as

$$\tilde{\gamma}_{p,k} = \frac{\text{Tr}(\bar{\mathbf{T}}_k \mathbf{f}_k \mathbf{f}_k^H)}{\text{Tr}(\bar{\mathbf{T}}_k (\mathbf{F}_a + \mathbf{F}_b)) - \text{Tr}(\bar{\mathbf{T}}_k \mathbf{f}_k \mathbf{f}_k^H) + (1 + \mu_r) \sigma_k^2}, \quad (17)$$

$$\tilde{\gamma}_{c,k} = \frac{\text{Tr}(\bar{\mathbf{T}}_k \mathbf{f}_c \mathbf{f}_c^H)}{\text{Tr}(\bar{\mathbf{T}}_k (\mathbf{F}_a + \mathbf{F}_b)) + (1 + \mu_r) \sigma_k^2}, \quad (18)$$

where $\mathbf{F}_a = \mu_r \mathbf{f}_c \mathbf{f}_c^H + (1 + \mu_r) \mu_t \text{diag}(\mathbf{f}_c \mathbf{f}_c^H)$ and $\mathbf{F}_b = (1 + \mu_r) (\mathbf{F} \mathbf{F}^H + \mu_t \text{diag}(\mathbf{F} \mathbf{F}^H))$, $\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_k] \in \mathbb{C}^{M \times K}$.

Furthermore, the average achievable private rate (AAPR) $R_{p,k}$ and the average achievable common rate (AACR) $R_{c,k}$ can be expressed as

$$\tilde{R}_{j,k} = \log_2(1 + \tilde{\gamma}_{j,k}), j = \{p, c\}. \quad (19)$$

Here, to make sure that all the users can successfully decode the common part of the message, the actual transmission rate for the common stream $\mathbf{c} = [c_1, c_2, \dots, c_K]$ should satisfy that $\sum_{k=1}^K c_k \leq \min_{k \in \mathcal{K}} R_{c,k}$, where c_k is the actual common rate allocated for the k -th user.

B. Problem Formulation

Considering the STAR-RIS aided downlink RSMA communication system with HWI, our objective is to maximize the overall achievable sum rate by jointly optimizing the beamforming vectors at the base station and STAR-RIS, as well as the actual transmission rate allocated to the common stream for each users. The optimization problem is formulated as follows:

$$(P_0) : \max_{\mathbf{f}_k, \mathbf{f}_c, \Phi_i, \mathbf{c}} \sum_{k=1}^K (\tilde{R}_{p,k} + c_k), \quad (20a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \|\mathbf{f}_k\|^2 + \|\mathbf{f}_c\|^2 \leq P_{max}, \quad (20b)$$

$$|\Phi_t|_{n,n}^2 + |\Phi_r|_{n,n}^2 = 1, \forall n, \quad (20c)$$

$$\sum_{k=1}^K c_k \leq \tilde{R}_{c,k}, \forall k, \quad (20d)$$

$$c_k \geq 0, \forall k. \quad (20e)$$

Constraint (20b) ensures that the transmit power remains within the allowable range of the maximum power budget P_{max} . Constraint (20c) indicates the amplitude of STAR-RIS elements should meet the law of energy conservation and the phase should satisfy the rules of the ES mode. Constraint (20d) guarantee the successfully decoding of the common stream, and (20e) illustrates that the actual achievable common rate should be non-negative.

III. PROPOSED ALGORITHM

In this section we propose a FP-AO algorithm to solve (P_0) . To deal with the highly complicated fractional term objective function, we first transform the original problem into a more tractable form based on the theory of FP. Then we alternatively optimize the auxiliary variables introduced by FP, the BS beamforming vector for both common and private stream as well as the actual common rate allocated for users, and the phase of the STAR-RIS.

A. Problem Reformulation by FP

The objective function of (P_0) is hard to tackle due to the sum of multiple logarithmic functions. Motivated by previous works [47], [48], [49], the closed-form FP techniques is used to simplify this sum-of-log-of-ratio problem. We first utilize lagrangian dual transform to deal with the log function by

introducing auxiliary variable $\mathbf{a} = [a_1, a_2, \dots, a_K]$, thereby (P_0) can be reformulated as

$$(P_1) : \max_{\mathbf{f}_k, \mathbf{f}_c, \Phi_i, \mathbf{a}, \mathbf{c}} g_R(\mathbf{f}_k, \mathbf{f}_c, \Phi_i, \mathbf{a}, \mathbf{c}), \quad (21)$$

$$\text{s.t.} \quad (20b), (20c), (20d), (20e),$$

where the objective function

$$g_R(\mathbf{f}_k, \mathbf{f}_c, \Phi_i, \mathbf{a}, \mathbf{c}) = \sum_{k=1}^K \left(\ln(1 + a_k) - a_k \right. \\ \left. + \frac{(1 + a_k) \text{Tr}(\bar{\mathbf{T}}_k \mathbf{f}_k \mathbf{f}_k^H)}{\text{Tr}(\bar{\mathbf{T}}_k (\mathbf{F}_a + \mathbf{F}_b)) + (1 + \mu_r) \sigma_k^2} \right) + \sum_{k=1}^K c_k. \quad (22)$$

However, the problem (P_1) is still hard to tackle since the presence of the sum of fractional terms. Upon applying quadratic transform [48] by introducing the auxiliary variable $\mathbf{b} = [b_1, b_2, \dots, b_K]$, the problem (P_1) can be further transformed into

$$(\widehat{P}_1) : \max_{\mathbf{f}_k, \mathbf{f}_c, \Phi_i, \mathbf{a}, \mathbf{b}, \mathbf{c}} f_R(\mathbf{f}_k, \mathbf{f}_c, \Phi_i, \mathbf{a}, \mathbf{b}, \mathbf{c}), \quad (24)$$

$$\text{s.t.} \quad (20b), (20c), (20d), (20e),$$

where $f_R(\mathbf{f}_k, \mathbf{f}_c, \Phi_i, \mathbf{a}, \mathbf{b}, \mathbf{c})$ is formulated as (23) at the top of the next page.

Compared with the original problem, the transferred form is more tractable by introducing two auxiliary variables. To solve $(\widehat{P}_1)$, we capitalize SCA and PF method alternatively to each group of variables while others kept fixed. The sub-problems of updating each group are specified in details in the following subsections.

B. Update the Auxiliary Variable Vectors

When variables $\mathbf{f}_k, \mathbf{f}_c, \Phi_i, \mathbf{c}$ are fixed, the auxiliary variable vectors $\mathbf{a}, \mathbf{b}$ introduced by FP will be easy to optimize by simply take partial derivative of (22) and (23) and follow the equation below:

$$\frac{\partial g_R(\mathbf{f}_k, \mathbf{f}_c, \Phi_i, \mathbf{a}, \mathbf{c})}{\partial a_k} = 0, \quad (25)$$

$$\frac{\partial f_R(\mathbf{f}_k, \mathbf{f}_c, \Phi_i, \mathbf{a}, \mathbf{b}, \mathbf{c})}{\partial b_k} = 0. \quad (26)$$

Therefore, the closed-form of the optimal a_k^*, b_k^* are given as

$$a_k^* = \tilde{\gamma}_{p,k}, \quad (27)$$

$$b_k^* = \frac{\sqrt{(1 + a_k) \text{Tr}(\bar{\mathbf{T}}_k \mathbf{f}_k \mathbf{f}_k^H)}}{\text{Tr}(\bar{\mathbf{T}}_k (\mathbf{F}_a + \mathbf{F}_b)) + (1 + \mu_r) \sigma_k^2}. \quad (28)$$

$$f_R(\mathbf{f}_k, \mathbf{f}_c, \Phi_i, \mathbf{a}, \mathbf{b}, \mathbf{c}) = \sum_{k=1}^K \left(\ln(1 + a_k) - a_k + 2b_k \sqrt{(1 + a_k) \text{Tr}(\bar{\mathbf{T}}_k \mathbf{f}_k \mathbf{f}_k^H)} - b_k^2 \left(\text{Tr}(\bar{\mathbf{T}}_k (\mathbf{F}_a + \mathbf{F}_b)) + (1 + \mu_r) \sigma_k^2 \right) + c_k \right), \quad (23)$$

C. Update the BS Beamforming and the Actual Common Rate

For fixed $\mathbf{a}, \mathbf{b}$ and Φ_i , we focus on the optimization of $\{\mathbf{f}_k, \mathbf{f}_c, \mathbf{c}\}$. By omitting the constant value in the objective function, the corresponding optimization problem can be rewritten as

$$(P_2) : \max_{\mathbf{f}_k, \mathbf{f}_c, \mathbf{c}} \sum_{k=1}^K \left(2b_k \sqrt{(1 + a_k) \text{Tr}(\bar{\mathbf{T}}_k \mathbf{f}_k \mathbf{f}_k^H)} - b_k^2 \text{Tr}(\bar{\mathbf{T}}_k (\mathbf{F}_a + \mathbf{F}_b)) + c_k \right), \quad (29)$$

s.t. (20b), (20d), (20e).

To further simplify the problem, let define $\tilde{\mathbf{F}}_k = \mathbf{f}_k \mathbf{f}_k^H$ and $\tilde{\mathbf{F}}_c = \mathbf{f}_c \mathbf{f}_c^H$, which should satisfy the constraints $\tilde{\mathbf{F}}_k \succeq \mathbf{0}$, $\tilde{\mathbf{F}}_c \succeq \mathbf{0}$ and $\text{rank}(\tilde{\mathbf{F}}_k) = \text{rank}(\tilde{\mathbf{F}}_c) = 1$. However, the rank-1 constraint is naturally satisfied here [50]. Hence, $\mathbf{F}_a$ and $\mathbf{F}_b$ can be rewritten as

$$\mathbf{F}'_a = \mu_r \tilde{\mathbf{F}}_c + (1 + \mu_r) \mu_t \text{diag}(\tilde{\mathbf{F}}_c), \quad (30)$$

$$\mathbf{F}'_b = (1 + \mu_r) \left(\sum_{k=1}^K \tilde{\mathbf{F}}_k + \mu_t \text{diag} \left(\sum_{k=1}^K \tilde{\mathbf{F}}_k \right) \right). \quad (31)$$

As a result, the objective function (29) can be rewritten as

$$f_{R2}(\tilde{\mathbf{F}}_k, \tilde{\mathbf{F}}_c, \mathbf{c}) = \sum_{k=1}^K \left(2b_k \sqrt{(1 + a_k) \text{Tr}(\bar{\mathbf{T}}_k \tilde{\mathbf{F}}_k)} - b_k^2 \text{Tr}(\bar{\mathbf{T}}_k (\mathbf{F}'_a + \mathbf{F}'_b)) + c_k \right), \quad (32)$$

and the AACR can be expressed as

$$\tilde{R}_{c,k} = \log_2 \left(\gamma_k \text{Tr}(\bar{\mathbf{T}}_k (\mathbf{F}'_a + \mathbf{F}'_b)) + \gamma_k \text{Tr}(\bar{\mathbf{T}}_k \tilde{\mathbf{F}}_c) + 1 \right) - \log_2 \left(\gamma_k \text{Tr}(\bar{\mathbf{T}}_k (\mathbf{F}'_a + \mathbf{F}'_b)) + 1 \right), \quad (33)$$

where we let $\gamma_k = \frac{1}{(1 + \mu_r) \sigma_k^2}$ for convenience. Hence, the problem is reformulated as

$$(\bar{P}_2) : \max_{\tilde{\mathbf{F}}_k, \tilde{\mathbf{F}}_c, \mathbf{c}} f_{R2}(\tilde{\mathbf{F}}_k, \tilde{\mathbf{F}}_c, \mathbf{c}), \quad (34a)$$

s.t. (20e),

$$\tilde{\mathbf{F}}_k \succeq \mathbf{0}, \tilde{\mathbf{F}}_c \succeq \mathbf{0}, \quad (34b)$$

$$\text{Tr} \left(\sum_{k=1}^K \tilde{\mathbf{F}}_k + \tilde{\mathbf{F}}_c \right) \leq P_{max}, \quad (34c)$$

$$\sum_{k=1}^K c_k \leq \tilde{R}_{c,k}, \forall k. \quad (34d)$$

However, the problem $(\bar{P}_2)$ is still hard to solve since the non-convex constraint (34d). According to [42], (34d) can be transformed into a tractable form by leveraging the SCA method.

Specifically, by adopting the first-order Taylor approximation, the right-hand-side of (34d) can be approximated as

$$\begin{aligned} \tilde{R}_{c,k} &\geq \log_2 \left(\gamma_k \text{Tr}(\bar{\mathbf{T}}_k (\mathbf{F}'_a + \mathbf{F}'_b)) + \gamma_k \text{Tr}(\bar{\mathbf{T}}_k \tilde{\mathbf{F}}_c) + 1 \right) \\ &\quad - \log_2 \left(\gamma_k \text{Tr}(\bar{\mathbf{T}}_k (\mathbf{F}'_a + \mathbf{F}'_b)) + 1 \right) \\ &\quad - \frac{\gamma_k \text{Tr}(\bar{\mathbf{T}}_k (\mathbf{F}'_a - \mathbf{F}'_a + \mathbf{F}'_b - \mathbf{F}'_b))}{\ln 2 (\gamma_k \text{Tr}(\bar{\mathbf{T}}_k (\mathbf{F}'_a + \mathbf{F}'_b)) + 1)} \triangleq \bar{R}_{c,k}, \end{aligned} \quad (35)$$

where $\mathbf{F}'_a$ and $\mathbf{F}'_b$ are the solutions obtained in the n -th

Algorithm 1 FP-AO Algorithm

Require: $\mathbf{f}_k, \mathbf{f}_c, \Phi_i, \mathbf{c}$

- 1: Use FP to transfer (20a) to (23) by introduce $\mathbf{a}$ and $\mathbf{b}$.
 - 2: **repeat**
 - 3: Update $\mathbf{a}, \mathbf{b}$.
 - 4: Update $\mathbf{f}_k, \mathbf{f}_c, \mathbf{c}$ by solving problem $\hat{P}_2$.
 - 5: Update Φ_i by solving problem $\hat{P}_3$.
 - 6: **until** The objective function (20a) convergence.
-

iteration. Hence, the problem $(\bar{P}_2)$ can be transformed into

$$(\hat{P}_2) : \max_{\tilde{\mathbf{F}}_k, \tilde{\mathbf{F}}_c, \mathbf{c}} f_{R2}(\tilde{\mathbf{F}}_k, \tilde{\mathbf{F}}_c, \mathbf{c}), \quad (36a)$$

s.t. (20e), (34b), (34c)

$$\sum_{k=1}^K c_k \leq \bar{R}_{c,k}, \forall k. \quad (36b)$$

At this stage, all the constraints are convex and the problem can be solved by CVX tool.

D. Update the Phase Matrix of STAR-RIS

For fixed $\mathbf{a}, \mathbf{b}, \mathbf{f}_c, \mathbf{f}_k$ and $\mathbf{c}$, we aim to optimize the phase shift matrix Φ_i of the STAR-RIS. By omitting the constant value in the objective function, the corresponding optimization problem can be expressed as

$$(P_3) : \max_{\Phi_i} \sum_{k=1}^K \left(2b_k \sqrt{(1 + a_k) \text{Tr}(\bar{\mathbf{T}}_k \mathbf{f}_k \mathbf{f}_k^H)} - b_k^2 \text{Tr}(\bar{\mathbf{T}}_k (\mathbf{F}_a + \mathbf{F}_b)) \right), \quad (37)$$

s.t. (20c), (20d).

By observing (37), we first extract and handle $\text{Tr}(\bar{\mathbf{T}}_k \mathbf{f}_k \mathbf{f}_k^H)$ and $\text{Tr}(\bar{\mathbf{T}}_k (\mathbf{F}_a + \mathbf{F}_b))$, as these are the only two parts containing the target variables to be optimized.

For $\bar{\mathbf{T}}_k$, let define $\bar{\mathbf{T}}_{k,t}$ and $\bar{\mathbf{T}}_{k,r}$ as the cases for $k \in \mathcal{T}$ and $k \in \mathcal{R}$, respectively. Considering the same form of these two variables, we first address $\bar{\mathbf{T}}_{k,t}$ by rewriting it as follows

$$\begin{aligned} \text{Tr}(\bar{\mathbf{T}}_{k,t} \mathbf{f}_k \mathbf{f}_k^H) &= \text{Tr}(\mathbf{G}^H \Phi_t^H \text{diag}(\mathbf{g}_k) \mathbf{V}_x \text{diag}(\mathbf{g}_k^H) \Phi_t \mathbf{G} \mathbf{f}_k \mathbf{f}_k^H) \\ &\quad + \frac{2}{\pi} \mathbf{f}_k^H \mathbf{G}^H \Phi_t^H \mathbf{g}_k \mathbf{h}_k^H \mathbf{f}_k + \frac{2}{\pi} \mathbf{f}_k^H \mathbf{h}_k \mathbf{g}_k^H \Phi_t \mathbf{G} \mathbf{f}_k \\ &\quad + \mathbf{f}_k^H \mathbf{h}_k \mathbf{h}_k^H \mathbf{f}_k, \end{aligned} \quad (38)$$

where

$$\mathbf{V}_x = \mathbf{I}_N + \mathbf{J}. \quad (39)$$

Then, each term in (38) is handled individually. For the first term we have

$$\text{Tr}(\mathbf{G}^H \Phi_t^H \text{diag}(\mathbf{g}_k) \mathbf{V}_x \text{diag}(\mathbf{g}_k^H) \Phi_t \mathbf{G} \mathbf{f}_k \mathbf{f}_k^H) = \mathbf{v}_t^H \mathbf{Q}_k \mathbf{v}_t, \quad (40)$$

where $\mathbf{Q}_k = (\text{diag}(\mathbf{g}_k) \mathbf{V}_x \text{diag}(\mathbf{g}_k^H)) \odot (\mathbf{G} \mathbf{f}_k \mathbf{f}_k^H \mathbf{G}^H)^T$. For the second term we have

$$\frac{2}{\pi} \mathbf{f}_k^H \mathbf{G}^H \Phi_t^H \mathbf{g}_k \mathbf{h}_k^H \mathbf{f}_k = \mathbf{v}_t^H \mathbf{q}_k, \quad (41)$$

where $\mathbf{q}_k = \frac{2}{\pi} \text{diag}(\mathbf{h}_k^H \mathbf{f}_k \mathbf{f}_k^H \mathbf{G}^H) \mathbf{g}_k$. The third term is equal to the conjugate transpose of the second which can be represented as

$$\frac{2}{\pi} \mathbf{f}_k^H \mathbf{h}_k \mathbf{g}_k^H \Phi_t \mathbf{G} \mathbf{f}_k = \mathbf{q}_k^H \mathbf{v}_t. \quad (42)$$

From above, the transformed term of (38) can be expressed as

$$\begin{aligned} \text{Tr}(\bar{\mathbf{T}}_{k,t} \mathbf{f}_k \mathbf{f}_k^H) &= \mathbf{v}_t^H \mathbf{Q}_k \mathbf{v}_t + \mathbf{v}_t^H \mathbf{q}_k + \mathbf{q}_k^H \mathbf{v}_t \\ &\quad + \mathbf{f}_k^H \mathbf{h}_k \mathbf{h}_k^H \mathbf{f}_k. \end{aligned} \quad (43)$$

To further transform the expression, we define two extended matrix

$$\bar{\mathbf{v}}_t = \begin{bmatrix} \mathbf{v}_t \\ 1 \end{bmatrix} = [\mathbf{v}_t^H \quad 1]^H, \quad (44)$$

and

$$\bar{\mathbf{Q}}_k = \begin{bmatrix} \mathbf{Q}_k & \mathbf{q}_k \\ \mathbf{q}_k^H & \mathbf{f}_k^H \mathbf{h}_k \mathbf{h}_k^H \mathbf{f}_k \end{bmatrix}. \quad (45)$$

Therefore,

$$\text{Tr}(\bar{\mathbf{T}}_{k,t} \mathbf{f}_k \mathbf{f}_k^H) = \bar{\mathbf{v}}_t^H \bar{\mathbf{Q}}_k \bar{\mathbf{v}}_t = \text{Tr}(\bar{\mathbf{Q}}_k \bar{\mathbf{V}}_t), \quad (46)$$

where $\bar{\mathbf{V}}_t = \bar{\mathbf{v}}_t \bar{\mathbf{v}}_t^H$.

Since $\text{Tr}(\bar{\mathbf{T}}_{k,t} \mathbf{f}_k \mathbf{f}_k^H)$, $\text{Tr}(\bar{\mathbf{T}}_{k,r} \mathbf{f}_k \mathbf{f}_k^H)$, $\text{Tr}(\bar{\mathbf{T}}_{k,t} (\mathbf{F}_a + \mathbf{F}_b))$ and $\text{Tr}(\bar{\mathbf{T}}_{k,r} (\mathbf{F}_a + \mathbf{F}_b))$ have similar forms, we can easily derive

$$\text{Tr}(\bar{\mathbf{T}}_{k,r} \mathbf{f}_k \mathbf{f}_k^H) = \text{Tr}(\bar{\mathbf{Q}}_k \bar{\mathbf{V}}_r), \quad (47)$$

where $\bar{\mathbf{V}}_r = \bar{\mathbf{v}}_r \bar{\mathbf{v}}_r^H$, and

$$\text{Tr}(\bar{\mathbf{T}}_{k,t} (\mathbf{F}_a + \mathbf{F}_b)) = \text{Tr}(\bar{\mathbf{S}}_k \bar{\mathbf{V}}_t), \quad (48)$$

$$\text{Tr}(\bar{\mathbf{T}}_{k,r} (\mathbf{F}_a + \mathbf{F}_b)) = \text{Tr}(\bar{\mathbf{S}}_k \bar{\mathbf{V}}_r), \quad (49)$$

where

$$\bar{\mathbf{S}}_k = \begin{bmatrix} \mathbf{S}_k & \mathbf{s}_k \\ \mathbf{s}_k^H & \mathbf{h}_k^H (\mathbf{F}_a + \mathbf{F}_b) \mathbf{h}_k \end{bmatrix}, \quad (50)$$

with

$$\mathbf{S}_k = (\text{diag}(\mathbf{g}_k) \mathbf{V}_x \text{diag}(\mathbf{g}_k^H)) \odot (\mathbf{G} (\mathbf{F}_a + \mathbf{F}_b) \mathbf{G}^H)^T, \quad (51)$$

$$\mathbf{s}_k = \frac{2}{\pi} \text{diag}(\mathbf{h}_k^H (\mathbf{F}_a + \mathbf{F}_b) \mathbf{G}^H) \mathbf{g}_k. \quad (52)$$

At this stage, the objective function (37) can be reformulated as

$$\begin{aligned} f_{R3}(\bar{\mathbf{V}}_t, \bar{\mathbf{V}}_r) &= \sum_{k=0}^{k_0} \left(2b_k \sqrt{(1+a_k) \text{Tr}(\bar{\mathbf{Q}}_k \bar{\mathbf{V}}_t)} - b_k^2 \text{Tr}(\bar{\mathbf{S}}_k \bar{\mathbf{V}}_t) \right) \\ &\quad + \sum_{k_0+1}^K \left(2b_k \sqrt{(1+a_k) \text{Tr}(\bar{\mathbf{Q}}_k \bar{\mathbf{V}}_r)} - b_k^2 \text{Tr}(\bar{\mathbf{S}}_k \bar{\mathbf{V}}_r) \right). \end{aligned} \quad (53)$$

And the problem (P_3) can be rewritten as

$$(\bar{P}_3) : \max_{\bar{\mathbf{V}}_t, \bar{\mathbf{V}}_r} f_{R3}(\bar{\mathbf{V}}_t, \bar{\mathbf{V}}_r), \quad (54a)$$

$$\text{s.t. } [\bar{\mathbf{V}}_t]_{n,n} + [\bar{\mathbf{V}}_r]_{n,n} = 1, n = 1, \dots, N, \quad (54b)$$

$$[\bar{\mathbf{V}}_t]_{N+1} = [\bar{\mathbf{V}}_r]_{N+1} = 1, \quad (54c)$$

$$\bar{\mathbf{V}}_t \succeq \mathbf{0}, \bar{\mathbf{V}}_r \succeq \mathbf{0}, \quad (54d)$$

$$\text{Rank}(\bar{\mathbf{V}}_t) = \text{Rank}(\bar{\mathbf{V}}_r) = 1, \quad (54e)$$

$$\sum_{k=1}^K c_k \leq \tilde{R}_{c,k}, \forall k. \quad (54f)$$

Here, the constraint (54f) can be expressed as

$$\begin{aligned} \left(\text{Tr}(\bar{\mathbf{S}}_k \bar{\mathbf{V}}_t) + \frac{1}{\gamma} \right) 2^{\sum_{k=1}^K c_k} &\leq \text{Tr}(\bar{\mathbf{S}}_k \bar{\mathbf{V}}_t) + \\ &\quad \text{Tr}(\bar{\mathbf{W}}_k \bar{\mathbf{V}}_t) + \frac{1}{\gamma}, k \in \mathcal{T}, \end{aligned} \quad (55)$$

$$\begin{aligned} \left(\text{Tr}(\bar{\mathbf{S}}_k \bar{\mathbf{V}}_r) + \frac{1}{\gamma} \right) 2^{\sum_{k=1}^K c_k} &\leq \text{Tr}(\bar{\mathbf{S}}_k \bar{\mathbf{V}}_r) + \\ &\quad \text{Tr}(\bar{\mathbf{W}}_k \bar{\mathbf{V}}_r) + \frac{1}{\gamma}, k \in \mathcal{R}, \end{aligned} \quad (56)$$

where

$$\bar{\mathbf{W}}_k = \begin{bmatrix} \mathbf{W}_k & \mathbf{w}_k \\ \mathbf{w}_k^H & \mathbf{f}_k^H \mathbf{h}_k \mathbf{h}_k^H \mathbf{f}_k \end{bmatrix}, \quad (57)$$

with

$$\mathbf{W}_k = (\text{diag}(\mathbf{g}_k) \mathbf{V}_x \text{diag}(\mathbf{g}_k^H)) \odot (\mathbf{G} \mathbf{f}_c \mathbf{f}_c^H \mathbf{G}^H)^T, \quad (58)$$

$$\mathbf{w}_k = \frac{2}{\pi} \text{diag}(\mathbf{h}_k^H \mathbf{f}_c \mathbf{f}_c^H \mathbf{G}^H) \mathbf{g}_k. \quad (59)$$

To this end, apart from the non-convex rank-one constraint (54e), all other constraints are convex. To solve it, we resort to the PF method to handle constraint (54e) [51]. Noticing the fact:

$$\text{Tr}(\bar{\mathbf{V}}_i) - \lambda_{\max}(\bar{\mathbf{V}}_i) = 0 \Rightarrow \text{Rank}(\bar{\mathbf{V}}_i) = 1, i = \{v, r\}, \quad (60)$$

where $\lambda_{\max}(\bar{\mathbf{V}}_i)$ denotes the largest eigenvalue of $\bar{\mathbf{V}}_i$, we incorporate the term $\text{Tr}(\bar{\mathbf{V}}_i) - \lambda_{\max}(\bar{\mathbf{V}}_i)$ into (54a) as the penalty term and reformulate the objective function as

$$f_{R3}(\bar{\mathbf{V}}_t, \bar{\mathbf{V}}_r) - \eta \sum_{i=\{v,r\}} \left(\text{Tr}(\bar{\mathbf{V}}_i) - \lambda_{\max}(\bar{\mathbf{V}}_i) \right), \quad (61)$$

where η is the penalty coefficient. Following the idea of PF, the new objective function aims to maximize $f_{R3}(\bar{\mathbf{V}}_t, \bar{\mathbf{V}}_r)$ and minimize the penalty term $\sum_{i=\{v,r\}} \left(\text{Tr}(\bar{\mathbf{V}}_i) - \lambda_{\max}(\bar{\mathbf{V}}_i) \right)$, simultaneously. When the penalty term approaches zero, $\bar{\mathbf{V}}_i$

can be considered to have only one non-zero eigenvalue, and thus the rank-one constraint (54e) is satisfied. Then, by applying SCA method, using the first-order Taylor expansion, the convex term $\lambda_{max}(\bar{\mathbf{V}}_i)$ should satisfy the following inequality

$$\lambda_{max}(\bar{\mathbf{V}}_i) \geq \lambda_{max}(\bar{\mathbf{V}}_i^n) + \text{Tr}\left(\iota_{max}(\bar{\mathbf{V}}_i^n) \iota_{max}^H(\bar{\mathbf{V}}_i^n)(\bar{\mathbf{V}}_i - \bar{\mathbf{V}}_i^n)\right), \quad (62)$$

where $\bar{\mathbf{V}}_i^n$ is a feasible point of $\bar{\mathbf{V}}_i$ and $\iota_{max}(\bar{\mathbf{V}}_i^n)$ represents the corresponding unit-norm eigenvector of $\lambda_{max}(\bar{\mathbf{V}}_i^n)$. Thus, the objective function can be further reformulated as

$$f'_{R3}(\bar{\mathbf{V}}_t, \bar{\mathbf{V}}_r) = f_{R3}(\bar{\mathbf{V}}_t, \bar{\mathbf{V}}_r) - \eta \sum_{i=\{v,r\}} \left(\text{Tr}(\bar{\mathbf{V}}_i) - \lambda_{max}(\bar{\mathbf{V}}_i^n) - \text{Tr}\left(\iota_{max}(\bar{\mathbf{V}}_i^n) \iota_{max}^H(\bar{\mathbf{V}}_i^n)(\bar{\mathbf{V}}_i - \bar{\mathbf{V}}_i^n)\right) \right). \quad (63)$$

At this stage, $(\bar{P}_3)$ can be rewritten as

$$\begin{aligned} (\bar{P}_3) : \max_{\bar{\mathbf{V}}_t, \bar{\mathbf{V}}_r} & f'_{R3}(\bar{\mathbf{V}}_t, \bar{\mathbf{V}}_r), \\ \text{s.t.} & (54b), (54c), (54d), (55), (56). \end{aligned} \quad (64a)$$

Till now, the problem can be solved by using the CVX tool.

E. Complexity and Optimality Analysis

1) *Complexity Analysis*: The computational complexity represents the number of multiplication operations in an expression. Generally, the highest-order term is taken while lower-order terms are ignored. Hence, for updating auxiliary variables a_k^* and b_k^* is $\mathcal{O}(2KM^3)$. For solving sub-problem (P_2) , we employ the SCA method, which contributes $\mathcal{O}(M^3)$ complexity per iteration. Assuming l_1 iterations for SCA, the total complexity of (P_2) becomes $\mathcal{O}(l_1(K+1)M^{3.5})$. Regarding sub-problem (P_3) , assume l_2 iterations for SCA, the computational complexity becomes $\mathcal{O}(2l_2(N+1)^{3.5})$.

2) *Optimality Analysis*: Based on the AO framework, the auxiliary variables a , b and sub-problems (P_2) , (P_3) are solved in an iterative way. For a and b , the optimal solution can be found by taking partial derivative while for (P_2) and (P_3) , both of which yield locally optimal solutions. Hence, the proposed FP-AO algorithm for solving (P_0) guarantees convergence to at least a locally optimal solution, as substantiated by the proof presented in [52].

IV. NUMERICAL RESULTS

In this section, associated numerical results are provided to evaluate the performance of the proposed STAR-RIS aided RSMA communication system and to validate the benefits of the robust beamforming design considering HWI.

A. Simulation Setup

In the simulations, we consider a single-cell situation that includes one BS, multiple users and one STAR-RIS which is installed close to the users. In a two-dimensional (2D) coordinate the BS is located at (0,0)m, STAR-RIS at (40,0)m, four users are introduced and each two of them are randomly distributed in two circular regions centered at (-2, 40)m and

TABLE I
SIMULATION PARAMETERS.

Parameters	Description	value
M	the amount of antennas	4
K	the amount of users	4
σ^2	noise power	-70 dBm [53]
ε	rician factor	10
L_0	path loss at 1m	-30 dB
c_G	path loss component of G	2.6
$c_{g,k}$	path loss component of g	2.2
c_h	path loss component of h	5
ϵ	convergence accuracy	10^{-4}

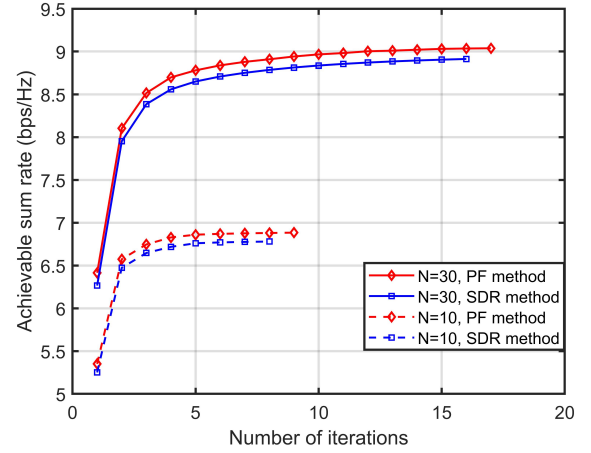


Fig. 2. Achievable sum rate versus the number of iterations to get convergence under different number of STAR-RIS elements N and different algorithms where $\mu_t = \mu_r = 0.01$, $P_{max} = 5$ W.

(2, 40)m with a radius 2 m, corresponding to the transmission region and reflection region respectively. The simulation parameters and descriptions are shown in Table I.

We consider that the links between the BS and the STAR-RIS, as well as those between the STAR-RIS and users, are characterized by LoS channels. Thus, both $\mathbf{G}$ and $\mathbf{g}_k$ are assumed to follow the Rician distribution.

$$\mathbf{G} = \beta_G \left(\sqrt{\frac{\varepsilon}{1+\varepsilon}} \mathbf{G}^{\text{LoS}} + \left(\sqrt{\frac{1}{1+\varepsilon}} \mathbf{G}^{\text{NLoS}} \right) \right), \quad (65)$$

$$\mathbf{g}_k = \beta_{g,k} \left(\sqrt{\frac{\varepsilon}{1+\varepsilon}} \mathbf{g}_k^{\text{LoS}} + \left(\sqrt{\frac{1}{1+\varepsilon}} \mathbf{g}_k^{\text{NLoS}} \right) \right), \quad (66)$$

ε represents the Rician factor, β_G and $\beta_{g,k}$ denote the pathloss from the BS to the STAR-RIS and from the STAR-RIS to the k -th user, respectively. $\mathbf{G}^{\text{LoS}}$ and $\mathbf{g}_k^{\text{LoS}}$ are the corresponding LoS components, while $\mathbf{G}^{\text{NLoS}}$ and $\mathbf{g}_k^{\text{NLoS}}$ are the non-LoS (NLoS) components [53]. Here, the pathloss β_G and $\beta_{g,k}$ are expressed as $\sqrt{L_0 d_G^{-c_G}}$, $\sqrt{L_0 d_{g,k}^{-c_{g,k}}}$ where d is the distance between the two objects. c represents the corresponding pathloss component, and L_0 is the pathloss at 1m. Assume the channel between BS to the users follow the same model but without LoS component.

We consider several schemes as baseline for comparison. (1) RSMA-STAR-nonRobust scheme utilizes RSMA and STAR-

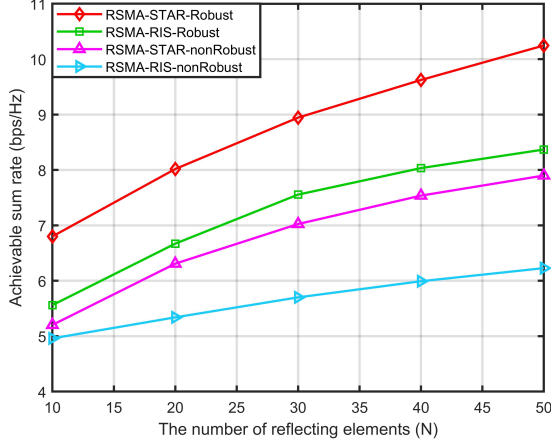


Fig. 3. Achievable sum rate versus the number of reflecting elements N under different type of RIS and beamforming design where $P_{max} = 5$ W, $\mu_t = \mu_r = 0.01$.

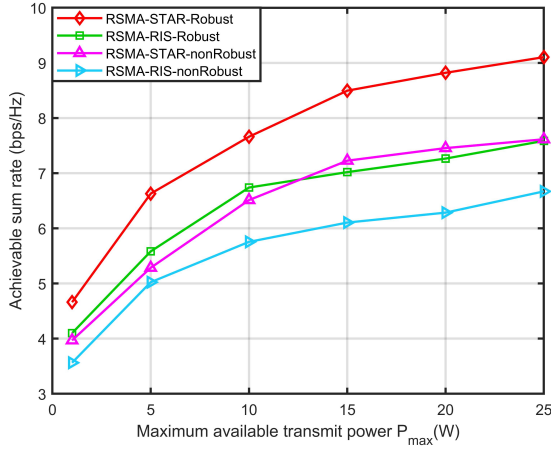


Fig. 4. Achievable sum rate versus the transmit power P_{max} under different type of RIS and beamforming design where $N = 10$, $\mu_t = \mu_r = 0.01$.

RIS in the system but neglects the influencing factors of HWI in the algorithm design; (2) RSMA-RIS-Robust scheme also employs RSMA method and considering the influence of HWI, while the STAR-RIS operates in a conventional RIS mode, where $N/2$ elements are used to transmit signals only, and $N/2$ elements are used to reflect signals only to facilitate full-space coverage; (3) RSMA-RIS-nonRobust scheme employs RSMA and conventional RIS mode, without considering HWI; (4) NOMA-STAR-Robust scheme adopts NOMA multiple-access technique aided by STAR-RIS while accounting for the presence of HWI where the deployment of NOMA scheme follows the one in [39]; (5) OMA-STAR-Robust scheme adopts OMA multiple-access technique aided by STAR-RIS while accounting for the presence of HWI. Here, we utilizes orthogonal frequency division multiple access (OFDMA) as the scheme of OMA which divided the frequency band equally into K parts for K users, respectively.

B. Simulation Result

1) *Convergence Performance*: Fig. 2 demonstrates that the proposed algorithm consistently converges to stable solutions through successive iterations, across varying numbers of STAR-RIS elements and users. As the number of STAR-RIS elements increases, it can be observed that the system's achievable sum rate exhibits improved performance. This enhancement is attributed to the fact that, under the same energy constraints, a larger number of elements offers greater degrees of freedom (DoF) for electromagnetic wave manipulation, leading to more precise beamforming and more flexible transmission strategy. However, it is evident that increasing the number of N also adds complexity to the problem, resulting in a higher number of iterations required for convergence. In the figure, it also compares the performance of different algorithm models by replacing the PF algorithm when solving the rank-1 constraint in $(\bar{P}_3)$ with the semidefinite relaxation (SDR) algorithm. The results demonstrate that our proposed algorithm outperforms the SDR-based approach. A possible explanation lies in the performance loss caused by the Gaussian randomization used in SDR.

2) *Gains of STAR-RIS and the Robust Beamforming Design*: To highlight the performance gains enabled by STAR-RIS and robust beamforming, Fig. 3 compares robust and non-robust designs in systems employing STAR-RIS and conventional passive RIS. It is observed that STAR-RIS consistently outperforms conventional RIS, regardless of whether HWI is taken into account. The primary reason for this improvement lies in the fact that STAR-RIS can fully exploit the available DoF to control beam propagation. In contrast, in conventional passive RIS-aided networks, the performance gain is limited due to the optimization being restricted solely to the reflection phases of the elements. Furthermore, when the same type of RIS is employed, the robust beamforming design consistently outperforms the non-robust approach. As the number of N increases, the performance advantage of the robust scheme over the non-robust scheme becomes increasingly evident. When $N = 10$, the achievable sum rate of RSMA-STAR-Robust scheme is approximately 21% higher than that of RSMA-STAR-nonRobust. When N reaches 50, this gap widens to 27.5%. This demonstrates the exceptional capability of the proposed FP-AO algorithm in compensating for HWI.

Fig. 4 plots the maximum system power budget versus the achievable sum rate. It shows that when P_{max} is small, the system achievable sum rate increases rapidly with the growth of the transmit power. As P_{max} rises from 1W to 5W, all four schemes achieve a significant performance improvement of approximately 40%. However, this trend gradually stabilizes as P_{max} increases. When P_{max} reaches 15W, further increasing the transmit power provides only a very marginal gain to the system. This stability appears because the distortion noise induced by HWI is proportional to the transceiver signal power, thereby establishing an upper bound for the performance of the system. Under the same power budget, it is evident that the robust design of the STAR-RIS aided RSMA system we proposed has better performance than the other three, demonstrates the necessity of considering HWI in the

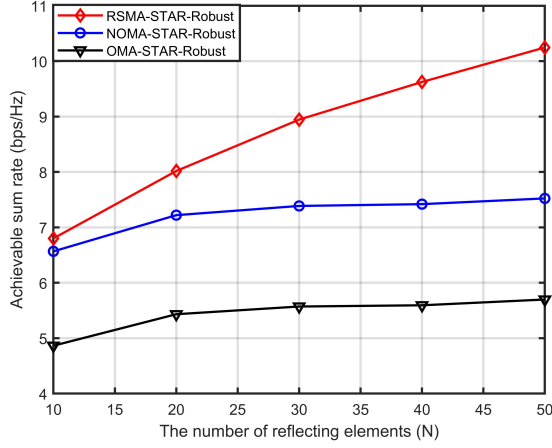


Fig. 5. Achievable sum rate versus the number of STAR-RIS elements N under different multiple access schemes where $P_{max} = 5$ W, $\mu_t = \mu_r = 0.01$.

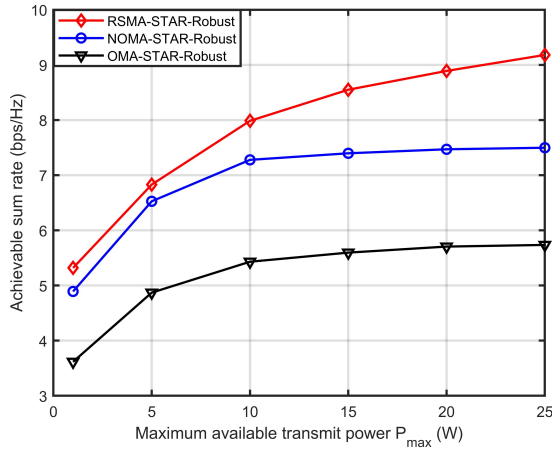


Fig. 6. Achievable sum rate versus the transmit power P_{max} under different multiple access schemes where $N = 10$, $\mu_t = \mu_r = 0.01$.

algorithm design.

3) *Effect of RSMA Scheme on the Network*: Fig. 5 presents the system achievable sum rate versus the number of reflecting elements N . As N increases, the performance of OMA and NOMA improves only marginally and even saturates when N becomes sufficiently large, due to the influence of HWI on reflecting elements. RSMA shows a more stable and superior trend which achieved nearly 50% performance gain as N increased from 10 to 50. At the point of $N = 50$, RSMA scheme delivers a 36% improvement over NOMA and a 75% advantage compared to traditional OMA scheme. This behavior suggests that RSMA is more resilient to the HWI on the reflecting elements, making it a more robust choice for realistic STAR-RIS systems. Fig. 6 compares the performance of different multiple access schemes in STAR-RIS Robust system. The results show that RSMA achieves the best performance across all transmit power levels, and its advantage becomes more significant as P_{max} increases. When power level is low, its increasing leads to substantial gains

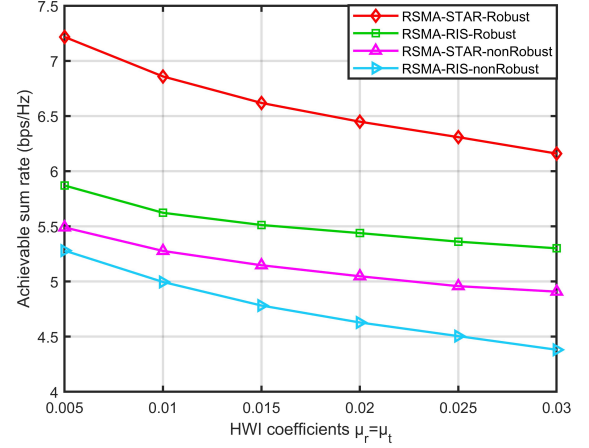


Fig. 7. Achievable sum rate versus the HWI coefficients μ_t, μ_r under different type of RIS and beamforming design where $N = 10$, $P_{max} = 5$ W.

in system performance. However, as P_{max} grows further, the performance improvements of all three schemes progressively decreases. This diminishing effect is likely caused by signal distortion under high-power conditions due to HWI's existing.

4) *Impact of the Transceiver HWI Coefficients*²: Fig. 7 shows the achievable sum rate of the system users under different transmission schemes versus the HWI coefficients. It is apparently noted that as the hardware impairments of the transceivers become more severe, the performance of each system declines accordingly, due to signal distortion caused by hardware degradation. It is observed that whether the robust design is implemented or not, the STAR-RIS aided RSMA system always performs better than the conventional RIS. Under low HWI conditions, the RSMA-STAR-Robust scheme achieves approximately a 23% gain in achievable sum rate compared to the RSMA-RIS-Robust scheme, while the RSMA-STAR-nonRobust scheme shows a 12% improvement over RSMA-RIS-nonRobust under high HWI conditions. Moreover, the robust design we proposed significantly outperforms non-robust designs under various levels of HWI impact and different type of RIS deployment, demonstrating the necessity of considering HWI in beamforming algorithms to overcome the impact of signal distortion interference and to better unlock the capabilities of RSMA scheme.

V. CONCLUSION

In this work, we explored a downlink RSMA communication system aided by a STAR-RIS, where the HWI is considered at both the transceivers and the STAR-RIS side. An achievable sum rate maximization problem is formulated by jointly optimizing the BS beamforming vectors for both the common and private streams, the STAR-RIS coefficients, and the actual transmission rate of the common stream. To tackle this highly coupled non-convexity problem, we first

²According to [41], the choice of HWI coefficients is closely related to the requirement of error vector magnitude (EVM) specified in the 3GPP standard. For NR and LTE systems, the HWI coefficients should be less than or equal to 0.175².

employ FP method to reformulate the objective function into a more tractable form, then solve the resulting sub-problems within an AO framework by leveraging SCA and PF methods. Simulation results show that the proposed FP-AO algorithm converges efficiently and that the RSMA-STAR-Robust scheme significantly outperforms benchmark schemes in terms of the achievable sum rate.

REFERENCES

- [1] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Netw.*, vol. 34, no. 3, pp. 134-142, May/Jun. 2020.
- [2] F. Tariq, M. R. A. Khandaker, K.-K. Wong, M. A. Imran, M. Bennis, and M. Debbah, "A speculative study on 6G," *IEEE Trans. Wireless Commun.*, vol. 27, no. 4, pp. 118-125, Aug. 2020.
- [3] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-lin, and Z. Wang, "Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74-81, Sept. 2015.
- [4] X. Foukas, G. Patounas, A. Elmokashfi, and M. K. Marina, "Network slicing in 5G: Survey and challenges," *IEEE Commun. Mag.*, vol. 55, no. 5, pp. 94-100, May 2017.
- [5] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, "A survey of non-orthogonal multiple access for 5G," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2294-2323, 3rd Quart., 2018.
- [6] Y. Liu, Z. Qin, M. El-kashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Nonorthogonal multiple access for 5G and beyond," *Proc. IEEE*, vol. 105, no. 12, pp. 2347-2381, Dec. 2017.
- [7] X. Guo, B. Li, J. Wu, R. Zhang, and X. Cheng, "Joint uplink and downlink NOMA for UAV relaying network with multi-pair users," *IEEE Trans. Wireless Commun.*, vol. 23, no. 12, pp. 18549-18562, Dec. 2024.
- [8] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501-1505, Dec. 2014.
- [9] Q. Wu, W. Chen, D. W. K. Ng, and R. Schober, "Spectral and energy-efficient wireless powered IoT networks: NOMA or TDMA?," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 6663-6667, July 2018.
- [10] B. Di, L. Song, and Y. Li, "Sub-channel assignment, power allocation, and user scheduling for non-orthogonal multiple access networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7686-7698, Nov. 2016.
- [11] B. Clerckx et al., "Is NOMA efficient in multi-antenna networks? A critical look at next generation multiple access techniques," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 1310-1343, 2021.
- [12] X. Lyu, S. Aditya, J. Kim, and B. Clerckx, "Rate-splitting multiple access: The first prototype and experimental validation of its superiority over SDMA and NOMA," *IEEE Trans. Wireless Commun.*, vol. 23, no. 8, pp. 9986-10000, Aug. 2024.
- [13] X. Liu, J. Feng, F. Li, and V. C. M. Leung, "Downlink energy efficiency maximization for RSMA-UAV assisted communications," *IEEE Wireless Commun. Lett.*, vol. 13, no. 1, pp. 98-102, Jan. 2024.
- [14] M. Can and I. Altunbas, "Channel-dependent power allocation and decoding order for RSMA-based uplink satellite networks," *IEEE Wireless Commun. Lett.*, vol. 13, no. 10, pp. 2822-2826, Oct. 2024.
- [15] J. Liu, Y. Liang Guan, Y. Ge, and B. Clerckx, "Power allocation for high mobility OTFS-RSMA system with path-selective one-tap receiver," *IEEE Wireless Commun. Lett.*, vol. 14, no. 3, pp. 661-665, March 2025.
- [16] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106-112, Jan. 2020.
- [17] X. Ma, S. Guo, H. Zhang, Y. Fang, and D. Yuan, "Joint beamforming and reflecting design in reconfigurable intelligent surface-aided multi-user communication systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 3269-3283, May 2021.
- [18] L. You, J. Xiong, D. W. K. Ng, C. Yuen, W. Wang, and X. Gao, "Energy efficiency and spectral efficiency tradeoff in RIS-aided multiuser MIMO uplink transmission," *IEEE Trans. Signal Process.*, vol. 69, pp. 1407-1421, 2021.
- [19] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157-4170, Aug. 2019.
- [20] H. Niu et al., "Joint beamforming design for secure RIS-assisted IoT networks," *IEEE Internet Things J.*, vol. 10, no. 2, pp. 1628-1641, Jan. 2023.
- [21] X. Peng, X. Hu, J. Gao, R. Jin, X. Chen, and C. Zhong, "Integrated localization and communication for IRS-assisted multi-user mmWave MIMO systems," *IEEE Trans. Commun.*, vol. 72, no. 8, pp. 4725-4740, Aug. 2024.
- [22] X. Li et al., "Covert Communications with Enhanced Physical Layer Security in RIS-Assisted Cooperative Networks," *IEEE Trans. Wireless Commun.*, doi: 10.1109/TWC.2025.3548024.
- [23] X. Li et al., "STAR-RIS-Assisted Covert Wireless Communications With Randomly Distributed Blockages," *IEEE Trans. Wireless Commun.*, vol. 24, no. 6, pp. 4690-4705, June 2025, doi: 10.1109/TWC.2025.3543421.
- [24] Z. Li, J. An, and C. Yuen, "Stacked intelligent metasurfaces for fully-analog wideband beamforming design," in *Proc. IEEE VTS APWCS*, Singapore, Aug. 2024, pp. 1-5.
- [25] Z. Li, J. An, and C. Yuen, "Stacked intelligent metasurfaces-enhanced MIMO OFDM wideband communication systems," *arXiv:2503.00368*, Mar. 2025.
- [26] T. Li, H. Zhang, S. Guo, and D. Yuan, "Max-min fair RIS-aided rate-splitting multiple access for multigroup multicast communications," *China Commun.*, vol. 20, no. 1, pp. 184-198, Jan. 2023.
- [27] B. Li et al., "Robust weighted sum-rate maximization for transmissive RIS transmitter enabled RSMA networks," *IEEE Commun. Lett.*, vol. 27, no. 10, pp. 2847-2851, Oct. 2023.
- [28] Y. Xu et al., "Robust beamforming and rate optimization for RIS-aided symbiotic radio systems with RSMA," *IEEE Commun. Lett.*, vol. 28, no. 10, pp. 2328-2332, Oct. 2024.
- [29] Y. Zhao et al., "2-bit RIS prototyping enhancing rapid-response space-time waveform manipulation for wireless communication: Experimental studies," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 4885-4901, 2024.
- [30] X. Mu, Y. Liu, L. Guo, J. Lin, and R. Schober, "Simultaneously transmitting and reflecting (STAR) RIS aided wireless communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 5, pp. 3083-3098, May 2022.
- [31] J. Zhang, Z. Li, L. Shao, and W. Zhu, "Dynamical absorption manipulation in a graphene-based optically transparent and flexible metasurface," *Carbon*, vol. 176, pp. 374-382, May 2021.
- [32] Y. Zhao et al., "2-bit intelligent reflection surface enhances urban wireless communications: An experimental study," in *Proc IEEE iWAT*, Sendai, Japan, 2024, pp. 186-189.
- [33] M. Ahmed et al., "A survey on STAR-RIS: Use cases, recent advances, and future research challenges," *IEEE Internet Things J.*, vol. 10, no. 16, pp. 14689-14711, Aug. 2023.
- [34] F. Karim, S. K. Singh, K. Singh, S. Prakriya, and C.-P. Li, "Performance analysis for RSMA-empowered STAR-RIS-aided downlink communications," in *Proc. IEEE PIMRC*, Toronto, ON, Canada, 2023, pp. 1-6.
- [35] C. Meng, C. Xiong, W. Chen, B. Gao, P. Fan, and K. B. Letaief, "Sum-rate maximization in STAR-RIS-assisted RSMA networks: A PPO-based algorithm," *IEEE Internet Things J.*, vol. 11, no. 4, pp. 5667-5680, Feb. 2024.
- [36] B. Wang, X. Tao, S. Han, K. Yang, and H. Wu, "Secure transmission design for rate-splitting empowered STAR-RIS-aided networks," *IEEE Wireless Commun. Lett.*, vol. 13, no. 9, pp. 2581-2585, Sept. 2024.
- [37] S. Ghosh, K. Singh, C. Pan, Q. Wu, and C.-P. Li, "On the performance analysis of RSMA based transmission in STAR-RIS-aided ISAC systems," in *Proc. IEEE ICC*, Denver, CO, USA, 2024, pp. 2470-2475.
- [38] Z. Peng, Z. Zhang, C. Pan, M. D. Renzo, O. A. Dobre, and J. Wang, "Beamforming optimization for active RIS-aided multiuser communications with hardware impairments," *IEEE Trans. Wireless Commun.*, vol. 23, no. 8, pp. 9884-9898, Aug. 2024.
- [39] Q. Li et al., "Achievable rate analysis of the STAR-RIS-aided NOMA uplink in the face of imperfect CSI and hardware impairments," *IEEE Trans. Commun.*, vol. 71, no. 10, pp. 6100-6114, Oct. 2023.
- [40] Z. Chu et al., "RIS assisted wireless powered IoT networks with phase shift error and transceiver hardware impairments," *IEEE Trans. Commun.*, vol. 70, no. 7, pp. 4910-4924, July 2022.
- [41] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7112-7139, Nov. 2014.
- [42] G. Zhou, C. Pan, H. Ren, K. Wang, and Z. Peng, "Secure wireless communication in RIS-aided MISO system with hardware impairments," *IEEE Wireless Commun. Lett.*, vol. 10, no. 6, pp. 1309-1313, June 2021.
- [43] C. Wu, C. You, Y. Liu, X. Gu and Y. Cai, "Channel Estimation for STAR-RIS-Aided Wireless Communication," *IEEE Wireless Commun. Lett.*, vol. 26, no. 3, pp. 652-656, March 2022.
- [44] M.-A. Badiu and J. P. Coon, "Communication through a large reflecting surface with phase errors," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 184-188, Feb. 2020.

- [45] Z. Xing, R. Wang, J. Wu, and E. Liu, "Achievable rate analysis and phase shift optimization on intelligent reflecting surface with hardware impairments," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 5514-5530, Sept. 2021.
- [46] Z. Peng, R. Weng, C. Pan, G. Zhou, M. Renzo, and A. Swindlehurst, "Robust transmission design for RIS-assisted secure multiuser communication systems in the presence of hardware impairments," *IEEE Trans. Wireless Commun.*, vol. 22, no. 11, pp. 7506-7521, Nov. 2023.
- [47] K. Shen and W. Yu, "Fractional programming for communication systems-Part I: Power control and beamforming," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2616-2630, May 2018.
- [48] K. Shen and W. Yu, "Fractional programming for communication systems-Part II: Uplink scheduling via matching," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2631-2644, May 2018.
- [49] Y. Ma, R. Liu, M. Li, Y. Liu, Q. Wu, and Q. Liu, "Reflection and relay dual-functional RIS assisted MU-MISO systems," in *Proc. IEEE WCNC*, Austin, TX, USA, 2022, pp. 387-392.
- [50] S. Zargari, A. Khalili, and R. Zhang, "Energy efficiency maximization via joint active and passive beamforming design for multiuser MISO IRS-aided SWIPT," *IEEE Wireless Commun. Lett.*, vol. 10, no. 3, pp. 557-561, Mar. 2021.
- [51] J. Ouyang et al., "Robust Secrecy-Energy Efficient Beamforming for Jittering UAV in Cognitive Satellite-Aerial Networks," *IEEE Trans. Aerosp. Electron. Syst.*, early access, 2025, doi: 10.1109/TAES.2025.3552313.
- [52] M. Hong et al., "A unified algorithmic framework for block-structured optimization involving Big Data: With applications in machine learning and signal processing," *IEEE Signal Process. Mag.*, vol. 33, no. 1, pp. 57-77, Jan. 2016.
- [53] Y. Li et al., "Robust and Secure Transmission Design for Reconfigurable Intelligent Surface Aided MU-SWIPT Systems With Hardware Impairments," *IEEE Trans. Veh. Technol.*, vol. 74, no. 3, pp. 4669-4683, March 2025.



Ziyue Wang received the B.Eng. degree in electronic and information from Beijing Jiaotong University in 2023, and the master's degree with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, in 2024.

His research interests include reconfigurable intelligent surface (RIS) and wireless signal processing.



Xiaoyan Ma received B.S. and Ph.D. degrees from the School of Information Science and Engineering, Shandong University, Jinan, China, in 2017 and 2023, respectively. From August 2023 to December 2024, she worked as a Research Fellow with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. In January 2025, she started to pursue her second PhD degree at Purdue University, West Lafayette, US. Her research interests include multiple-in-multiple-out technology, reconfigurable intelligent surface, orbital angular momentum waveform design, and federated learning.



Xingyu Peng received the B.S. degree in information science and engineering from Shandong University, Qingdao, China, in 2021. He is currently pursuing the Ph.D. degree with Institute of Information and Communication Engineering, Zhejiang University. His research interests include intelligent reflecting surface and integrated sensing and communication.



Zheao Li (Graduate Student Member, IEEE) received the B.E. and M.E. degrees in information engineering in Communication Engineering from the School of Information Science and Engineering, Southeast University, Nanjing, China, in 2020 and 2023, respectively. He is currently pursuing the Ph.D. degree with the School of Electrical and Electronic Engineering, Nanyang Technological University (NTU), Singapore. His current research interests include stacked intelligent metasurfaces (SIM), reconfigurable intelligent surfaces (RIS), wideband

MIMO-OFDM communication systems, and predictive channel modeling. His work focuses on wave-domain signal processing and optimization algorithms for enhancing the spectral and energy efficiency of next-generation wireless networks.



Jinyuan Liu (Graduate Student Member, IEEE) received the B.Eng. degree from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2020, and the M.Sc. degree from Imperial College London, London, U.K., in 2021. He is currently pursuing the Ph.D. degree with Nanyang Technological University, Singapore. His current research interests include rate-splitting multiple access (RSMA), fluid antenna systems (FAS), and advanced waveforms for next-generation wireless communication networks.



Yongliang Guan (Senior Member, IEEE) obtained his PhD degree from the Imperial College London, UK, and Bachelor of Engineering degree with first class honours from the National University of Singapore. He is now an Associate Vice President of the Nanyang Technological University (NTU), Singapore, and a Professor of Communication Engineering at the School of Electrical and Electronic Engineering in NTU, where he founded and is leading the Continental-NTU Corporate Research Lab, and also led the successful deployment of the

campus-wide NTU-NXP V2X Test Bed in NTU. His research interests broadly include coding and signal processing for communication systems and data storage systems. He has published an invited monograph, 2 books and more than 540 journal and conference papers. He has secured over S\$90 million of external research funding. He has more than 40 filed patents and 13 granted patents (some of which were licensed to NXP, Continental). He is an Editor for the IEEE Transactions on Vehicular Technology and a Distinguished Speaker of the IEEE Vehicular Technology Society.



Chau Yuen (S02-M06-SM12-F21) received the B.Eng. and Ph.D. degrees from Nanyang Technological University, Singapore, in 2000 and 2004, respectively. He was a Post-Doctoral Fellow with Lucent Technologies Bell Labs, Murray Hill, in 2005. From 2006 to 2010, he was with the Institute for Infocomm Research, Singapore. From 2010 to 2023, he was with the Engineering Product Development Pillar, Singapore University of Technology and Design. Since 2023, he has been with the School of Electrical and Electronic Engineering, Nanyang Technological

University, currently he is Provost's Chair in Wireless Communications, Assistant Dean in Graduate College, and Cluster Director for Sustainable Built Environment at ER@IN.

Dr. Yuen received IEEE Communications Society Leonard G. Abraham Prize (2024), IEEE Communications Society Best Tutorial Paper Award (2024), IEEE Communications Society Fred W. Ellersick Prize (2023), IEEE Marconi Prize Paper Award in Wireless Communications (2021), IEEE APB Outstanding Paper Award (2023), and EURASIP Best Paper Award for JOURNAL ON WIRELESS COMMUNICATIONS AND NETWORKING (2021).

Dr Yuen current serves as an Editor-in-Chief for Springer Nature Computer Science, Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, and IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING, where he was awarded as IEEE TNSE Excellent Editor Award 2024 and 2022, and Top Associate Editor for TVT from 2009 to 2015. He also served as the guest editor for several special issues, including IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE WIRELESS COMMUNICATIONS MAGAZINE, IEEE COMMUNICATIONS MAGAZINE, IEEE VEHICULAR TECHNOLOGY MAGAZINE, IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING, and ELSEVIER APPLIED ENERGY.

He is listed as Top 2% Scientists by Stanford University, and also a Highly Cited Researcher by Clarivate Web of Science from 2022.