Energy Efficiency in RSMA-Enhanced Active RIS-Aided Quantized Downlink Systems

Thanh Phung Truong, Thi My Tuyen Nguyen, The Vi Nguyen, Nhu-Ngoc Dao, and Sungrae Cho

***Abstract*—This work explores combining the rate-splitting multiple-access (RSMA) technique with an active reconfigurable intelligent surface (RIS) to improve the quantized multiuser multiple-input single-output network. The active RIS facilitates communication between the base station (BS) and users equipped with low-resolution quantizers, whereas RSMA improves downlink transmission efficiency. By maximizing the spectral efficiency while minimizing the power consumption at the transmitter and active RIS, we formulate an energy efficiency maximization problem by jointly designing the BS precoding matrix and active RIS reflecting matrix. The optimization problem presents non-convexity, which makes finding the optimal solution challenging. Therefore, we reformulate the problem into a reinforcement learning-based problem that is solvable by applying deep reinforcement learning (DRL) algorithms. To ensure action accuracy, we design a constraint-matching function that integrates with the DRL algorithm, forming a DRL framework securing all problem constraints. To assess the proposed DRL algorithm, we propose an alternating-based solution that decomposes the problem into precoding matrix optimization and active reflecting matrix optimization sub-problems, which are solvable using the successive convex approximation-based method. The performance evaluations demonstrate the convergence and effectiveness of the proposed approaches in various scenarios.**

***Index Terms*—Active reconfigurable intelligent surface, low-resolution quantizers, rate-splitting multiple access**

I. Introduction

THE demand for efficient transmission has become more significant in the burgeoning wireless communication technology era. However, traditional communication systems often require help operating in scenarios with constrained resources or stringent hardware limitations. Therefore, implementing low-resolution quantizers at transceivers has been examined in many studies to reduce the energy consumption in multiantenna systems [1]–[3]. Low-resolution quantizers lower the power usage at transceivers by reducing the quantization levels in digital-to-analog converters (DACs) and analog-to-digital converters (ADCs) [4], [5]. Although these quantizers effectively reduce power consumption, they concurrently generate signal distortion, causing quantization errors in transmissions. This error may cause interference and noise within the signal, leading to a reduction in transmission efficiency. Regarding this limitation, the quest to refine and optimize quantized systems has emerged as a prospect for exploration, especially considering the evolving technology to enhance the efficiency of quantized systems.

As an innovative technology for efficient transmission, the reconfigurable intelligent surface (RIS) has recently been widely researched. Traditionally, RIS operates as a passive structure comprising an array of programmable reflecting elements, enabling precise control over the phase of reflected signals and effectively shaping the electromagnetic wave to achieve the desired outcomes [6], [7]. However, recent research advancements have proposed the active RIS, integrating reflection-type amplifiers to adjust the phase and amplitude characteristics, demonstrating its effectiveness compared to the passive RIS [8]–[10]. Although active RIS has improved transmission efficiency in many studies [11]–[14], these often assume perfect quantization, overlooking quantized distortion. This omission makes performance characterization in low-resolution systems unpredictable. Therefore, exploring RIS-aided systems with low-resolution quantizers presents a suitable approach to tackling the difficulties of energy-efficient devices in communication systems.

Besides, the proliferation of wireless devices has intensified the demand for advanced multiple-access techniques to enhance communication performance. In this context, rate-splitting multiple access (RSMA) [15], [16] has been regarded as a potential and effective multiple-access scheme for the next generation of communication [17], [18]. Many studies have proved RSMA's effectiveness in enhancing communication efficiency compared with other multiple-access techniques [19]–[22], where it can improve throughput, lower latency, and enhance the computing system performance. Given these developments, integrating RSMA into RIS-aided MISO systems with low-resolution quantizers presents an intriguing research opportunity. This combination has the potential to address multiple challenges simultaneously: i) the need for efficient multiple access, ii) the benefits of RIS-aided transmission, and iii) the practical scenario of low-resolution hardware for energy-efficient transmission. Exploring this synergy between RSMA, RIS, and low-resolution quantizers could open new avenues for advancing next-generation wireless communication systems, particularly in scenarios where hardware limitations and increasing user density are significant concerns.

Motivated by these observations, this work explores the potential of the combination of the active RIS and RSMA in enhancing the energy efficiency (EE) of quantized multiuser

This work was supported in part by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. RS-2024-00453301), and in part by the Chung-Ang University Young Scientist Scholarship in 2023.

Corresponding authors: Nhu-Ngoc Dao and Sungrae Cho.

Thanh Phung Truong, Thi My Tuyen Nguyen, The Vi Nguyen, and Sungrae Cho are with the School of Computer Science and Engineering, Chung-Ang University, Seoul 06974, Republic of Korea (e-mail: {tptruong, tuyen, tvnguyen}@uclab.re.kr; srcho@cau.ac.kr).

Nhu-Ngoc Dao is with the Department of Computer Science and Engineering, Sejong University, Seoul 05006, Republic of Korea (e-mail: nndao@sejong.ac.kr).

multiple-input single-output (MISO) systems. The major contributions are summarized as follows.

- We study a novel quantized downlink multiuser MISO system that combines the active RIS and RSMA for efficient transmission. The active RIS facilitates communication between the base station (BS) and users obstructed by obstacles, and RSMA is employed to improve downlink transmission communication. In such a system, we formulate an EE maximization problem to maximize spectral efficiency (SE) while minimizing the transmitter and active RIS power consumption by optimizing the BS precoding matrix and active RIS reflecting matrix.
- As the problem exposes nonconvexity, we propose a deep reinforcement learning (DRL) algorithm to solve it. Here, we transform the problem into a reinforcement learning-based problem, which is solvable by applying DRL algorithms. To secure the action constraints, we propose a constraint-matching function to integrate with the DRL algorithm, forming a DRL framework that efficiently designs the BS precoding and active RIS reflecting matrices while ensuring action accuracy.
- To assess the proposed DRL solution, we design a successive convex approximation (SCA)-based alternating optimization approach, decomposing the problem into two sub-problems: precoding and active reflecting matrix optimizations. Accordingly, we apply several mathematical methods to approximate each sub-problem into a resolvable convex problem using convex optimization tools.
- We establish several numerical simulations to demonstrate the effectiveness of the proposed approaches. We evaluate the convergence of the DRL-based and alternating solutions by observing the change in system performance. In addition, we demonstrate the superior performance of the proposed system and algorithms compared with other benchmark schemes under various environmental scenarios. Additionally, we analyze the influence of the quantization levels on the transmission efficiency by varying the resolution at the transceivers.

The remaining sections of this study are organized as follows. Section II summarizes the related work. Section III introduces the proposed system . Then, Section IV formulates the EE problem. Accordingly, Sections V and VI present the DRL-based and SCA-based alternating optimization solutions, respectively, and Section VII describes the numerical results. Finally, Section VIII concludes the work.

II. Related Work

With the demand for efficient transmission while maintaining low power consumption, low-resolution converters have recently been considered in the literature [2], [23], [24]. For instance, the authors in [2] considered the EE maximization problem in a massive multiple-input multiple-output (MIMO) downlink system, where transceivers are equipped with low-resolution converters. The EE maximization problem was also investigated in [23], where the authors proposed a quantized hybrid precoding problem. Using an alternating optimization approach, they jointly designed the analog precoding and digital baseband precoding matrices to maximize the achievable instantaneous rate while considering the power consumption at the transmitter. In [24], a mixed-ADC/DAC architecture was investigated to reduce the power consumption in a MIMO system. A power allocation problem was formulated, and the SE and EE performance results were analyzed in detail. Moreover , recent advancements in research have prompted a closer examination of evolving technologies to enhance the efficacy of quantized network systems, especially with RIS [25], [26] and multiple access techniques [27]–[29]. In particular, the authors in [25] discussed the assistance of RIS in a massive multiuser MIMO system, where the BS is equipped with low-resolution DACs. They optimized the RIS phase shift to maximize the system achievable rate. Cascaded channel estimation was examined for a millimeter-wave MIMO system in [26], where the RIS aids communication between a user and a BS equipped with low-resolution ADCs. In this work, the authors proposed a bilinear generalized approximate message passing (BiG-AMP)-based algorithm to estimate the cascaded channel under loss of information due to quantization. The application of non-orthogonal multiple access (NOMA) in low-resolution ADCs/DACs was examined in [27], exploring the SE and EE performance by applying closed-form expressions. The consideration of NOMA in low-resolution ADC systems was also explored in [28]. By applying the first-order Karush–Kuhn–Tucker (KKT) condition, the authors proposed a precoding algorithm to maximize the sum rate of the system. The advantages of RSMA were examined in quantized systems. For example, Park *et al.* [29] studied the improvement of RSMA in a multiuser MIMO system. They proposed a generalized power iteration algorithm to optimize the RSMA precoding matrix to maximize SE and demonstrate the effectiveness of RSMA in quantized systems, especially its superior performance compared with other multiple-access techniques. Despite these advancements, the potential synergy between RIS and advanced multiple access techniques in the context of quantized systems still needs to be explored.

Meanwhile, the effectiveness of the combination of RIS and RSMA has been demonstrated in prior studies [30]–[32]. For example, the RIS-aided uplink RSMA system was studied in [30]. To maximize the system sum rate, the author proposed an alternating-based algorithm using the SCA and Riemannian conjugate gradient algorithms to optimize the passive phase-shift matrix, decoding order, and users' power allocation. The results demonstrate the effectiveness of combining RIS and RSMA and highlight the improved performance of RSMA compared with other multiple-access schemes. The EE maximization problem in simultaneous wireless information and power transfer with passive RIS and RSMA was explored in [31]. To solve the problem, the authors proposed two approximation approaches, DRL-based and SCA-based solutions, to design the passive RIS phase shift, message rates, power-splitting ratios, and beamforming vectors at the transmitter. The combination of active RIS and RSMA was examined in [32]. This paper investigated the SE-EE trade-off and proposes an alternating optimization algorithm for designing BS precoding and active RIS reflecting matrices. In addition,

TABLE I: Comparison of the proposed study with related works

Research	Low-resolution quantizers	RIS	Multiple access	Objective	Optimization m ethod
[2]	✓	✗	✗	Energy efficiency	Alternating algorithm
[23]	✓	✗	✗	Energy efficiency	Alternating algorithm
[24]	✓	✗	✗	Spectral/energy efficiency	SCA-based solution
[25]	✓	Passive RIS	✗	Achievable system rate	Particle swarm optimization
[26]	✓	Passive RIS	✗	Estimated cascaded channel	BiG-AMP-based algorithm
[27]	✓	✗	NOMA	Energy/spectral efficiency	Closed-form expressions
[28]	✓	✗	NOMA	System sum rate	KKT-based solution
[29]	✓	✗	RSMA	Spectral efficiency	Generalized iteration-based solution
[30]	✗	Passive RIS	RSMA	System sum rate	Alternating-based solution
[31]	✗	Passive RIS	RSMA	Energy efficiency	DRL-based and SCA-based solutions
[32]	✗	Active RIS	RSMA	Energy/spectral efficiency trade-off	Alternating optimization
Proposed	✓	Active RIS	RSMA	Energy efficiency	Alternating-based and DRL-based solutions

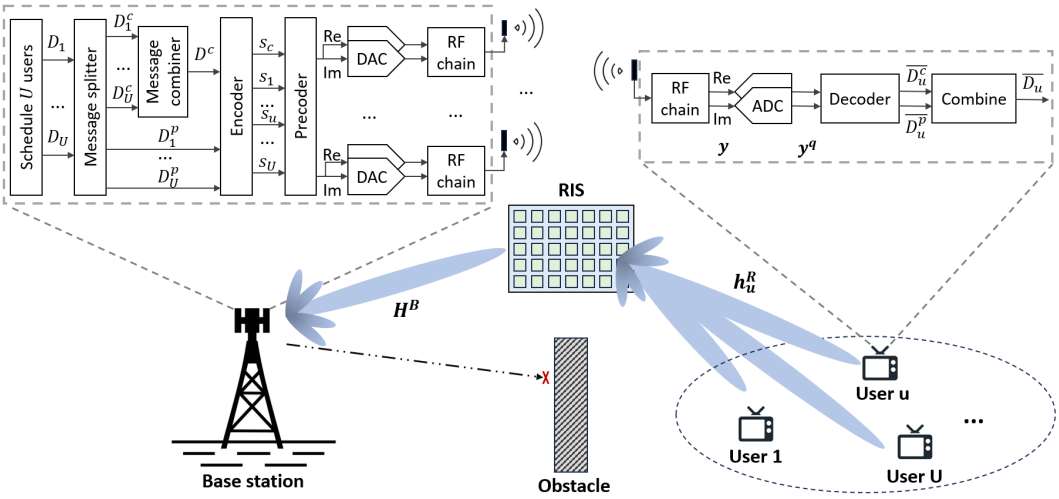


Fig. 1: RSMA-enhanced RIS-aided quantized downlink multiuser MISO system.

the simulation results proved the superiority of the active RIS compared with the passive RIS when integrating with RSMA. Therefore, applying this combination underscores the potential enhancements to the system, particularly in quantized systems. Integrating RIS and RSMA in low-resolution quantized systems could address the current energy efficiency and spectral utilization limitations, especially in scenarios where hardware constraints and spectral efficiency are critical considerations.

Table I summarizes the related work to highlight the novelty and contributions of this work.

III. System Model

We investigate a downlink multiuser MISO system in which an M-antenna BS serves U single-antenna users via RIS with N phase-shift elements, as illustrated in Fig. 1. We build the system by inheriting from previous studies [17], [29], [33], where obstacles block the direct links from BS to users[1]; hence, the BS transmits signals to users with the aid of the RIS, and the downlink transmission is enhanced by applying the one -layer RSMA technique. At the BS, each antenna is deployed with a b_m^{DAC}-bit resolution DAC pair (for imaginary and real parts), where b_m^{DAC} denotes the number of quantization bits (NoQBs) at antenna m. Similarly, at the receiver side, each user u is deployed with a b_u^{ADC}-bit resolution ADC pair (for imaginary and real parts), where b_u^{ADC} is the NoQBs at user u. The description of the entire system can be divided into three stages: *1) quantized transmitter*, where the BS generates the transmit signal by applying the RSMA technique with DACs to transmit to users; *2) transmission with RIS*, the communication between the BS and users via RIS; and *3) quantized receivers*, where the received signal is converted via ADCs and decoded according to the RSMA technique.

[1]As discussed in many previous works [30], [34]–[36], the RIS is effectively applied in environments where direct links between transmitters and receivers do not exist due to the unfavorable propagation and the sophistication of urban environments.

A. Quantized Transmitter

Following the one-layer RSMA, the BS splits each transmit message D_u (intended for u-th users) into a common submessage D_u^c and a private submessage D_u^p. Then, it combines all common submessages D_u^c, $u \in \mathcal{U} \triangleq \{1, 2, \ldots, U\}$, into one common message D^c. Accordingly, an encoder is applied to encode the common message D^c into a common stream, s_c, and independently encode the private submessages into private streams, s_u, $u \in \mathcal{U}$. Thus, U transmit messages generate $U+1$ transmit streams, combining the common and private streams. By defining $\mathbf{s} = [s_c, s_1, ..., s_U]^T \in \mathbb{C}^{(U+1)\times 1}$, without loss of generality, we assume that $\mathbb{E}\{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}_{U+1}$. The streams are precoded by a linear precoder with a precoding matrix $\mathbf{W} \triangleq [\mathbf{w}_c, \mathbf{w}_1, \ldots, \mathbf{w}_U] \in \mathbb{C}^{M\times(U+1)}$, where $\mathbf{w}_c, \mathbf{w}_u \in \mathbb{C}^{M\times 1}$ are the corresponding precoding vectors. Consequently, the digital

TABLE II: Quantization distortion factor

NoQBs	1	2	3	4	5
QDF	0.3634	0.1175	0.03454	0.009497	0.002499

transmit signal is given as

$$\mathbf{x} = \mathbf{W}\mathbf{s} = \mathbf{w}_c s_c + \sum_{u=1}^{U} \mathbf{w}_u s_u. \tag{1}$$

Accordingly, the digital signal is quantized with the DACs, where the quantization process follows the AQNM method [37], using a linear form to approximate the quantization process. By denoting $\mathcal{Q}(.)$ a quantizer function, the quantized signal, $\mathbf{x^q} \in \mathbb{C}^{M\times 1}$, is expressed as [29]

$$\mathbf{x^q} = \mathcal{Q}(\mathbf{x}) \approx \mathbf{\Theta}_\theta^{\mathbf{DAC}} \mathbf{w}_c s_c + \mathbf{\Theta}_\theta^{\mathbf{DAC}} \sum_{u=1}^{U} \mathbf{w}_u s_u + \mathbf{e^{DAC}}, \tag{2}$$

where $\mathbf{e^{DAC}} \in \mathbb{C}^{M\times 1}$ is the additive Gaussian quantization noise (AGQN) vector at the DACs, and $\mathbf{\Theta}_\theta^{\mathbf{DAC}} \triangleq \text{diag}(\theta_1^{DAC}, \ldots, \theta_M^{DAC})$ denotes the quantization loss matrix. Here, $\theta_m^{DAC} \in (0,1)$, $m \in \mathcal{M} \triangleq \{1,2,\ldots,M\}$, is the quantization loss of the m-th element calculated according to the quantization distortion factor (QDF), ϑ_m^{DAC}, determined as

$$\theta_m^{DAC} = 1 - \vartheta_m^{DAC}, \tag{3}$$

where ϑ_m^{DAC} is specified in Table II if $b_m^{DAC} \leq 5$; otherwise, $\vartheta_m^{DAC} = \frac{\pi\sqrt{3}}{2} 2^{-2b_m^{DAC}}$ [23]. The AGQN follows $\mathbf{e^{DAC}} \sim \mathcal{CN}(\mathbf{0_{M\times 1}}, \mathbf{R^{DAC}})$, where $\mathbf{R^{DAC}}$ is the covariance matrix of $\mathbf{e^{DAC}}$, computed as

$$\mathbf{R}^{\mathbf{DAC}} = \mathbf{\Theta}_\theta^{\mathbf{DAC}} \mathbf{\Theta}_\vartheta^{\mathbf{DAC}} \text{diag}\left(\mathbb{E}\left[\mathbf{x}\mathbf{x}^H\right]\right), \tag{4}$$

where $\mathbf{\Theta}_\vartheta^{\mathbf{DAC}} \triangleq \text{diag}\left(\vartheta_1^{DAC}, \ldots, \vartheta_M^{DAC}\right)$. Let P_{Bmax} denote the maximum transmission power at the BS, the quantized transmit signal has a power constraint as [38]

$$\text{tr}\left(\mathbb{E}\left[\mathbf{x^q}(\mathbf{x^q})^H\right]\right) \leq P_{Bmax}. \tag{5}$$

Proposition 1. Given the real diagonal matrix $\mathbf{\Theta}_\theta^{DAC} \triangleq \text{diag}(\theta_1^{DAC}, \theta_2^{DAC}, \ldots, \theta_M^{DAC})$, the transmit power constraint in (5) can be rewritten as follows

$$\text{tr}\left(\mathbf{\Theta}_\theta^{\mathbf{DAC}} \mathbf{W}\mathbf{W}^H\right) \leq P_{Bmax}. \tag{6}$$

Proof. Please see Appendix A. □

B. Transmission with the active RIS

This study considers that the direct links from the BS to the users are neglected due to obstacles and unfavorable propagation conditions. Therefore, the BS transmits its signal to users via the RIS link. Denoting $\mathbf{H^B} \in \mathbb{C}^{N\times M}$ and $\mathbf{h_u^R} \in \mathbb{C}^{1\times N}$ are the channel matrices from the BS to RIS and from the RIS to user u, respectively, the received signal at u-th user is expressed as

$$y_u = \mathbf{h_u^R} \mathbf{\Phi} \mathbf{H^B} \mathbf{x^q} + \mathbf{h_u^R} \mathbf{\Phi} \omega_r + \omega_u, \tag{7}$$

where $\mathbf{\Phi} \triangleq \text{diag}\left(\phi_1, \ldots, \phi_N\right)$, $\phi_n \triangleq \psi_n e^{j\varphi_n}, \psi_n \in [0, \psi_{max}], \varphi_n \in [0, 2\pi], n \in \mathcal{N} \triangleq \{1,2,\ldots,N\}$, denotes the active reflecting matrix at the RIS applied to the reflection [39], $\omega_r \sim \mathcal{CN}(\mathbf{0}_N, \sigma_r^2 \mathbf{1}_N)$ denotes the thermal noise generated at the active RIS, and ω_u is the additive white Gaussian noise (AWGN) for user u with a zero mean and variance of σ_u^2.

C. Quantized Receivers

The received signal y_u for user u is quantized by b_u^{ADC}-bit resolution ADCs. Then, the digital signal received by the u-th user is expressed as

$$y_u^q = \mathcal{Q}(y_u) \approx \theta_u^{ADC} y_u + e_u^{ADC}, \tag{8}$$

where e_u^{ADC} is the AGQN value, and $\theta_u^{ADC} \triangleq 1 - \vartheta_u^{ADC}$ denotes the quantization loss with ϑ_u^{ADC} as the QDF of the ADCs for the u-th user. Similarly, the value of ϑ_u^{ADC} is specified in Table II if $b_u^{ADC} \leq 5$; otherwise, $\vartheta_u^{ADC} = \frac{\pi\sqrt{3}}{2} 2^{-2b_u^{ADC}}$. Based on [37], the value of e_u^{ADC} follows $\mathcal{CN}(0, r_u^{ADC})$, where r_u^{ADC} is calculated as

$$r_u^{ADC} = \theta_u^{ADC} \vartheta_u^{ADC} \mathbb{E}(|y_u|^2). \tag{9}$$

Accordingly, user u decodes the received streams using the SIC technique, where the common stream is decoded into a common message, $\bar{D}^c$, by treating all private streams as interference, and the private stream of user u is decoded into its private message, $\bar{D}_u^p$, by treating the remaining private streams as interference. Then, user u extracts its common message $\bar{D}_u^c$ from $\bar{D}^c$ and combines it with $\bar{D}_u^p$ to reconstruct the original message $\bar{D}_u$.

IV. Problem Statement

A. Performance Metrics

1) Spectral Efficiency: The digital signal received by user u can be decomposed into four components: common stream, private stream, interference from other users' private streams, and noise from transmission and quantization. From (2), (7), and (8), the received digital signal for the u-th user is represented as

$$\begin{aligned} y_u^q =& \theta_u^{ADC} \mathbf{h_u^R} \mathbf{\Phi} \mathbf{H^B} \left(\mathbf{\Theta}_\theta^{\mathbf{DAC}} \mathbf{w}_c s_c + \mathbf{\Theta}_\theta^{\mathbf{DAC}} \sum_{u=1}^{U} \mathbf{w}_u s_u + \mathbf{e^{DAC}}\right) \\ &+ \theta_u^{ADC} \mathbf{h_u^R} \mathbf{\Phi} \omega_r + \theta_u^{ADC} \omega_u + e_u^{ADC} \\ =& \underbrace{\theta_u^{ADC} \mathbf{h_u^R} \mathbf{\Phi} \mathbf{H^B} \mathbf{\Theta}_\theta^{\mathbf{DAC}} \mathbf{w}_c s_c}_{\text{common stream}} + \underbrace{\theta_u^{ADC} \mathbf{h_u^R} \mathbf{\Phi} \mathbf{H^B} \mathbf{\Theta}_\theta^{\mathbf{DAC}} \mathbf{w}_u s_u}_{\text{private stream}} \\ &+ \underbrace{\theta_u^{ADC} \mathbf{h_u^R} \mathbf{\Phi} \mathbf{H^B} \mathbf{\Theta}_\theta^{\mathbf{DAC}} \sum_{k\in\{\mathcal{U}\setminus u\}} \mathbf{w}_k s_k}_{\text{interference}} \\ &+ \underbrace{\theta_u^{ADC} \mathbf{h_u^R} \mathbf{\Phi} \omega_r + \theta_u^{ADC} \mathbf{h_u^R} \mathbf{\Phi} \mathbf{H^B} \mathbf{e^{DAC}} + \theta_u^{ADC} \omega_u + e_u^{ADC}}_{\text{noise}}. \end{aligned} \tag{10}$$

Accordingly, the achievable transmission rate of the common stream for user u is determined as follows:

$$r_u^c = \log_2\left(1 + \frac{|\theta_u^{ADC} \mathbf{h_u^R} \mathbf{\Phi} \mathbf{H^B} \mathbf{\Theta}_\theta^{\mathbf{DAC}} \mathbf{w}_c|^2}{\sum_{k=1}^{U} |\theta_u^{ADC} \mathbf{h_u^R} \mathbf{\Phi} \mathbf{H^B} \mathbf{\Theta}_\theta^{\mathbf{DAC}} \mathbf{w}_k|^2 + \text{NE}_u}\right), \tag{11}$$

where $\text{NE}_u = (\theta_u^{ADC})^2 \mathbf{h}_u^{\mathbf{R}} \mathbf{\Phi} (\mathbf{h}_u^{\mathbf{R}} \mathbf{\Phi})^H \sigma_r^2 + (\theta_u^{ADC})^2 \mathbf{h}_u^{\mathbf{R}} \mathbf{\Phi} \mathbf{H}^{\mathbf{B}} \mathbf{R}^{\mathbf{DAC}} (\mathbf{h}_u^{\mathbf{R}} \mathbf{\Phi} \mathbf{H}^{\mathbf{B}})^H + (\theta_u^{ADC})^2 \sigma_u^2 + r_u^{ADC}$. To guarantee that all users successfully decode the common stream, the common stream's rate, r^c, is determined as

$$r^c = \min\{r_1^c, r_2^c, \ldots, r_U^c\}. \tag{12}$$

The common stream is combined and encoded from the common submessages of all users; therefore, the transmission rate of the common submessage (common rate) for each user is a portion of r^c. Consequently, by denoting c_u the common rate allocated to user u, the transmission rates of common submessages have to satisfy the following constraint

$$\sum_{u \in \mathcal{U}} c_u \leq r^c. \tag{13}$$

After successfully decoding and subtracting the common stream from the received signal, user u decodes its private stream from the remaining signals. Then, the achievable transmission rate of the private stream (private rate) of user u is determined as follows:

$$r_u^p = \log_2 \left(1 + \frac{|\theta_u^{ADC} \mathbf{h}_u^{\mathbf{R}} \mathbf{\Phi} \mathbf{H}^{\mathbf{B}} \mathbf{\Theta}_\theta^{\mathbf{DAC}} \mathbf{w}_u|^2}{\text{IN}_u + \text{NE}_u} \right), \tag{14}$$

where $\text{IN}_u = \sum_{k \in \{\mathcal{U} \setminus u\}} |\theta_u^{ADC} \mathbf{h}_u^{\mathbf{R}} \mathbf{\Phi} \mathbf{H}^{\mathbf{B}} \mathbf{\Theta}_\theta^{\mathbf{DAC}} \mathbf{w}_k|^2$. As a result, the achievable rate of user u is determined as

$$r_u = c_u + r_u^p. \tag{15}$$

2) Transmitter Power Consumption: The power consumption at the BS comprise power amplifier (P_{PA}) and analog circuits (P_{AC}) [2], [23]. The P_{PA} value is calculated based on the transmit power (P_T), expressed as

$$P_{PA} = \eta^{-1} P_T, \tag{16}$$

where η denotes the power-added efficiency. As defined in sub-section III-A, the transmit power is calculated as

$$P_T = \text{tr}\big(\mathbb{E}\big[\mathbf{x}^{\mathbf{q}} (\mathbf{x}^{\mathbf{q}})^H\big]\big). \tag{17}$$

Meanwhile, the power consumption of the analog circuits comprises the power consumption of the local oscillator (P_{LO}), DACs ($P_{DAC,m}, m \in \mathcal{M}$), and radio frequency (RF) chain (P_{RF}), calculated as [2]

$$P_{AC} = P_{LO} + \sum_{m \in \mathcal{M}} (2 P_{DAC,m} + P_{RF}). \tag{18}$$

According to [23], the DACs' power consumption of m-th antenna is calculated as

$$P_{DAC,m} = 1.5 \times 10^{-5} \cdot 2^{b_m^{DAC}} + 9 \times 10^{-12} \cdot b_m^{DAC} \cdot F_s, \tag{19}$$

where F_s denotes the sampling frequency [40]. The value of P_{RF} is determined as

$$P_{RF} = 2 P_{LP} + 2 P_M + P_H, \tag{20}$$

where P_{LP}, P_M, and P_H denote the power consumption of the low-pass filter, mixer, and 90° hybrid with buffer, respectively. Consequently, the BS power consumption is calculated as

$$P_{BS} = P_{AC} + P_{PA}. \tag{21}$$

To clarify, Table III summarizes the power-related parameters and their values, which are compiled from [2], [23].

TABLE III: Power-related parameters.

Parameters	Values
Power-added efficiency, η	27%
Local oscillator power consumption, P_{LO}	22.5 mW
Sampling rate, F_s	1 GHz
Low-pass filter power consumption, P_{LP}	14 mW
Mixer power consumption, P_M	0.3 mW
90° hybrid with buffer power consumption, P_H	3 mW

3) RIS Power Consumption: Let $\mathbf{y_r}$ denote the signal reflected and amplified by the active RIS, it can be expressed as

$$\mathbf{y_r} = \mathbf{\Phi} \mathbf{H}^{\mathbf{B}} \mathbf{x}^{\mathbf{q}} + \mathbf{\Phi} \omega_r. \tag{22}$$

According to [41], the power consumption of the active RIS can be computed as

$$P_R = N(P_c + P_{DC}) + \xi^{-1} P_{R-out}, \tag{23}$$

where P_c and P_{DC} denote the switch and control circuit and the DC biasing power consumption for each reflecting element, P_{R-out} indicates the output power of the RIS, and ξ represents the amplifier efficiency coefficient. The RIS output power is

$$\begin{aligned} P_{R-out} &= \mathbf{tr}(\mathbb{E}[\| \mathbf{y_r} \|^2]) \\ &= \mathbf{tr}\left(\mathbf{\Phi} \mathbf{H}^{\mathbf{B}} \mathbb{E}[\mathbf{x}^{\mathbf{q}} (\mathbf{x}^{\mathbf{q}})^H] (\mathbf{\Phi} \mathbf{H}^{\mathbf{B}})^H + \mathbf{\Phi} \mathbf{\Phi}^H \sigma_r^2 \right) \\ &\overset{(a)}{=} \text{tr}\left(\mathbf{\Phi} \mathbf{H}^{\mathbf{B}} \mathbf{\Theta}_\theta^{\mathbf{DAC}} \mathbf{W} \mathbf{W}^H (\mathbf{\Phi} \mathbf{H}^{\mathbf{B}})^H + \mathbf{\Phi} \mathbf{\Phi}^H \sigma_r^2 \right), \end{aligned} \tag{24}$$

where (a) is obtained from Appendix A. The active RIS power consumption should follow the following power constraint:

$$P_R \leq P_{Rmax}, \tag{25}$$

where P_{Rmax} is the maximum amplification power of the active RIS.

4) Energy Efficiency: We establish the EE according to the sum of the users' SE and the system power consumption, including the BS and active RIS power consumption. As described in sub-section IV-A1, the sum of users' SE is determined as

$$SE = \sum_{u \in \mathcal{U}} r_u = \sum_{u \in \mathcal{U}} (c_u + r_u^p). \tag{26}$$

Remark 1. Since the target is to maximize the SE, we design the sum of common rates as the maximum value, i.e., the constraint in (13) becomes $\sum_{u \in \mathcal{U}} c_u = r^c$. Then, the distribution of the common rate for each user does not affect the primary goal. Accordingly, the SE is rewritten as follows:

$$SE = r^c + \sum_{u \in \mathcal{U}} r_u^p. \tag{27}$$

Consequently, the EE is calculated by dividing the SE by the system power consumption, expressed as

$$EE = \frac{SE}{P_{BS} + P_R}. \tag{28}$$

B. Problem Formulation

To handle the multiple-target objective of maximizing the SE while minimizing the power consumption of the transmitter and active RIS, we formulate an optimization problem of maximizing the EE by jointly designing the precoding matrix and active RIS reflecting matrix. Accordingly, the formulation of the problem is expressed as

$$\begin{aligned}
\text{(P1):}\quad \max_{\mathbf{W},\mathbf{\Phi}}\quad & EE && \text{(29a)}\\
\text{s.t.}\quad & \text{tr}\Big(\mathbf{\Theta}_{\theta}^{\mathbf{DAC}}\mathbf{W}\mathbf{W}^{H}\Big) \le P_{Bmax}, && \text{(29b)}\\
& P_R \le P_{Rmax}, && \text{(29c)}\\
& |\phi_n| \le \psi_{max}, n \in \mathcal{N}, && \text{(29d)}
\end{aligned}$$

where (29b) and (29c) are the constraints of power consumption at the BS and RIS, respectively, and (29d) indicates the value range of the RIS elements.

Solving Problem (P1) is challenging because the objective is nonconvex, and optimization variables are highly coupled. Fortunately, as demonstrated in numerous recent works [42]–[45], DRL has emerged as an efficient tool for addressing such problems. Therefore, we propose a DRL-based method to solve the considered problem. However, a key issue is how to evaluate the proposed DRL method. To address this, we propose another approximation method that alternately optimizes the variables based on the successive convex approximation (SCA) method. This method is an effective solution to such a problem, as evidenced by its performance in numerous prior studies [30]–[32]. Accordingly, we propose two approximate approaches: first, we transform the problem into a reinforcement learning-based problem and propose a DRL framework to resolve it; second, we relax the nonconvex objective function and apply an alternating optimization method.

V. Deep Reinforcement Learning-based Solution

A. Reinforcement Learning-based Problem

To solve the nonconvex problem (P1), we propose a DRL framework named *Proposed-DRL*, which is an effective approach for solving such a problem. First, we transform (P1) into an RL-based problem, which contains an agent trained to decide actions based on the environment's state to get a suitable reward. With high-power computation complexity, the agent is deployed at the BS, and the entire system is the working environment. Here, the state space, action space, and reward function are defined as follows:

- *State space:* At each time step, the state space indicates the change in the environment. In this system, the state space contains the channel condition between the BS and users obtained by transmitting via the RIS links [46]. Accordingly, the state space at time slot t is expressed as
$$s[t] = \{\mathbf{h}_u^{\mathbf{R}}[t], \mathbf{H}^{\mathbf{B}}[t], u \in \mathcal{U}\}. \tag{30}$$
- *Action space:* The action space contains all the variables the agent has to decide. Therefore, the action space in this system contains the precoding matrix and the active RIS reflecting matrix. At time slot t, it is expressed as
$$a[t] = \{\mathbf{W}[t], \mathbf{\Phi}[t]\}. \tag{31}$$
- *Reward function:* The reward function indicates how well the action is designed. In this maximization problem, the reward is calculated as the objective function, which is calculated as
$$r[t] = \frac{r^c[t] + \sum_{u\in\mathcal{C}} r_u^p[t]}{P_R[t] + P_{BS}[t]}. \tag{32}$$

The agent reinforces its decision through the training process by interacting with the environment and updating the policy according to the received reward. To update the policy, many DRL algorithms have been proposed and demonstrated their effectiveness in AI-related research. In this work, we apply a well-known DRL algorithm named deep deterministic policy gradient (DDPG) [47], which designs actions using neural networks (NNs). Many previous studies have demonstrated its efficacy [42]–[45].

B. Proposed Deep Reinforcement Learning framework

1) Constraint-matching Function: By applying the training algorithm, the agent is trained and updated on its policy to decide the appropriate action. However, the action decided by the NNs may not satisfy the constraints in (29), violating the system's essence. To address this challenge, let $\mathbf{W}^{\mathbf{d}} \triangleq [\mathbf{w}_c^{\mathbf{d}}, \mathbf{w}_1^{\mathbf{d}}, \ldots, \mathbf{w}_U^{\mathbf{d}}]$ and $\mathbf{\Phi}^{\mathbf{d}} \triangleq \text{diag}\big(\phi_1^d, \ldots, \phi_N^d\big)$ denote the precoding and active reflecting matrices obtained by the NNs, respectively, we propose a constraint-matching function that maps $\mathbf{W}^{\mathbf{d}}$ and $\mathbf{\Phi}^{\mathbf{d}}$ to the corresponding $\mathbf{W}$ and $\mathbf{\Phi}$, satisfying the constraints. This function is denoted as $\mathcal{M}(\cdot)$.

To ensure constraint (29b), we introduce a new action, $p_B \in [0, 1]$, denoting the fraction of power the BS uses for transmission. Let $\mathbf{W}' \triangleq \frac{1}{\sqrt{p_B P_{Bmax}}}\mathbf{W}$, i.e., $\mathbf{W} = \mathbf{W}'\sqrt{p_B P_{Bmax}}$, denote the normalized precoding matrix so that we rewrite the constraint (29b) as $p_B P_{Bmax}\text{tr}\Big(\mathbf{\Theta}_{\theta}^{\text{DAC}}\mathbf{W}'(\mathbf{W}')^H\Big) \le P_{Bmax}$. Accordingly, constraint (29b) can be equivalently split into two sub-constraints as

$$\begin{aligned}
& p_B \in [0, 1], && \text{(33a)}\\
& \text{tr}\Big(\mathbf{\Theta}_{\theta}^{\text{DAC}}\mathbf{W}'(\mathbf{W}')^H\Big) = 1. && \text{(33b)}
\end{aligned}$$

By using the DRL algorithm that applies NNs to decide action, we can normalize the range of the action by the activation functions. According to the constraint (33a), we normalize the decided action to the range of $[0, 1]$, satisfying this constraint. To ensure constraint (33b), we propose the following proposition.

Proposition 2. To satisfy constraint (33b), element at m-th row and u-th column of $\mathbf{W}'$, $w'_{m,u}$, can be calculated as

$$w'_{m,u} = \frac{w_{m,u}^d}{\sqrt{\sum_{m=1}^{M}\theta_m^{DAC}\sum_{u=1}^{U+1}|w_{m,u}^d|^2}}, \tag{34}$$

where $w_{m,u}^d$ is the element at m-th row and u-th column of $\mathbf{W}^{\mathbf{d}}$.

Proof. Based on (34), we aim to prove that the normalized precoding matrix, $\mathbf{W}'$, satisfies constraint (33b). Then, we first express the multiplication inside the trace operation, which is performed in (35) at the top of the next page. Accordingly,

$$\begin{aligned}\boldsymbol{\Theta}_{\theta}^{\mathbf{DAC}}\mathbf{W}'(\mathbf{W}')^{H} &= \operatorname{diag}\Big(\theta_1^{DAC},\ldots,\theta_M^{DAC}\Big)\begin{bmatrix} w'_{1,1} & w'_{1,2} & \cdots & w'_{1,U+1}\\ \cdots & \cdots & \cdots & \cdots \\ w'_{M,1} & w'_{M,2} & \cdots & w'_{M,U+1}\end{bmatrix}\begin{bmatrix} w'_{1,1} & w'_{1,2} & \cdots & w'_{1,U+1}\\ \cdots & \cdots & \cdots & \cdots \\ w'_{M,1} & w'_{M,2} & \cdots & w'_{M,U+1}\end{bmatrix}^{H}\\ &= \begin{bmatrix}\theta_1^{DAC}\sum_{u=1}^{U+1}|w'_{1,u}|^2 & \cdots & \cdots & \cdots\\ \cdots & \theta_2^{DAC}\sum_{u=1}^{U+1}|w'_{2,u}|^2 & \cdots & \cdots\\ \cdots & \cdots & \cdots & \theta_M^{DAC}\sum_{u=1}^{U+1}|w'_{M,u}|^2\end{bmatrix}.\end{aligned} \tag{35}$$

the left-hand side of the constraint (33b) can be calculated as

$$\begin{aligned}\operatorname{tr}\Big(\boldsymbol{\Theta}_{\theta}^{\mathrm{DAC}}\mathbf{W}'(\mathbf{W}')^{H}\Big) &= \sum_{m=1}^{M}\theta_m^{DAC}\sum_{u=1}^{U+1}|w'_{m,u}|^2\\ &= \sum_{m=1}^{M}\theta_m^{DAC}\sum_{u=1}^{U+1}\frac{|w^d_{m,u}|^2}{\sum_{m=1}^{M}\theta_m^{DAC}\sum_{u=1}^{U+1}|w^d_{m,u}|^2}\\ &= 1.\end{aligned} \tag{36}$$

This completes the proof. □

Next, we let $\boldsymbol{\Phi}' \triangleq \operatorname{diag}\big(\phi'_1,\ldots,\phi'_N\big)$ denote the normalized active reflecting matrix that satisfies the constraint (29d). With the action range of $[0,1]$, i.e. $\psi_n^d, \varphi_n^d \in [0,1]$, the n-th elements in $\boldsymbol{\Phi}'$, $\phi'_n \triangleq \psi'_n e^{j\varphi'_n}$, can be calculated as

$$\begin{aligned}\psi'_n &= \psi_n^d * \psi_{max},\\ \varphi'_n &= \varphi_n^d * 2\pi.\end{aligned} \tag{37}$$

To ensure constraint (29c), we introduce a new action, $p_r \in [0,1]$, as the fraction of power the RIS uses. Then, the power consumption at the RIS can be presented as

$$P_R = p_r P_{Rmax}. \tag{38}$$

By substituting (23) and (24) into (38) , constraint (29c) can be rewritten as

$$\begin{aligned}&\mathbf{tr}\Big(\boldsymbol{\Phi}\mathbf{H}^{\mathbf{B}}\boldsymbol{\Theta}_{\theta}^{\mathbf{DAC}}\mathbf{W}\mathbf{W}^{H}(\boldsymbol{\Phi}\mathbf{H}^{\mathbf{B}})^{H} + \boldsymbol{\Phi}\boldsymbol{\Phi}^{H}\sigma_r^2\Big)\\ &= \xi\left(p_r P_{Rmax} - N(P_c + P_{DC})\right).\end{aligned} \tag{39}$$

As the active reflecting matrix $\boldsymbol{\Phi}'$ may not satisfy the above constraint, we normalize it using the following proposition to ensure that (39) holds at every time slot.

Proposition 3. By denoting $\mathbf{G} \triangleq \mathbf{H}^{\mathbf{B}}\boldsymbol{\Theta}_{\theta}^{\mathbf{DAC}\,\frac{1}{2}}\mathbf{W}$, constraint (29c) can be satisfied by normalizing the active reflecting matrix as follows

$$\boldsymbol{\Phi} = \begin{cases}\boldsymbol{\Phi}', & \text{if } \boldsymbol{\Phi}' \text{ satisfies (29c)},\\ C\boldsymbol{\Phi}', & \text{if } \boldsymbol{\Phi}' \text{ violates (29c)},\end{cases} \tag{40}$$

where $C \triangleq \dfrac{\sqrt{\xi\Big(p_r P_{Rmax} - N(P_c+P_{DC})\Big)}}{\sqrt{\sum_{n=1}^{N}|\phi'_n|^2\Big(\sum_{u=1}^{U+1}|g_{n,u}|^2+\sigma_r^2\Big)}}$, $g_{n,u}$ is the element at n-th row and u-th column of $\mathbf{G}$.

Proof. Please see Appendix B. □

By applying Proposition 3, constraint (29c) is satisfied. It can be verified that the reflecting matrix $\boldsymbol{\Phi}$ obtained from (40) also satisfies (29d). Particularly, if $\boldsymbol{\Phi}'$ satisfies (29c), then $\boldsymbol{\Phi} = \boldsymbol{\Phi}'$, which immediately follows that $\boldsymbol{\Phi}$ satisfies (29d). Otherwise, if $\boldsymbol{\Phi}'$ violates (29c), then $\boldsymbol{\Phi} = C\boldsymbol{\Phi}'$. We can verify that the normalizing factor C is less than 1, which is proved by employing the following lemma.

Algorithm 1 *Proposed-DRL* algorithm

1: Set up algorithm parameters.
2: **while** $e < E$ **do**
3: **for** t from 1 to T **do**
4: Observe state $s[t]$.
5: Generalize action with actor network:
$$a^d[t] \triangleq \{p_B[t], p_r[t], \mathbf{W}^{\mathbf{d}}[t], \boldsymbol{\Phi}^{\mathbf{d}}[t]\} = \mu^{\theta_\mu}(s[t]).$$
6: Recalculate action: $\mathbf{W}[t], \boldsymbol{\Phi}[t] = \mathcal{M}(a^d[t])$:
 ∘ Calculate $\mathbf{W}'[t]$ as (34) ⇒ $\mathbf{W}[t] = \mathbf{W}'[t]\sqrt{p_B[t]P_{Bmax}}$,
 ∘ Calculate $\boldsymbol{\Phi}'[t]$ as (37) ⇒ Calculate $\boldsymbol{\Phi}[t]$ as (40).
7: Perform $\mathbf{W}[t], \boldsymbol{\Phi}[t]$, observe next state $s'[t]$, reward $r[t]$.
8: Store $(s[t], a^d[t], s'[t], r[t])$ in buffer.
9: Update new state $s'[t] \to s[t]$.
10: Randomly choose training samples from the buffer.
11: Update NNs as (47), (48), (49).
12: **end for**
13: **end while**
14: **return** the trained main actor network, $\mu^{\theta^*_\mu}$.

Lemma 1. If $\boldsymbol{\Phi}'$ violates (29c), the following inequality holds

$$\frac{\sqrt{\xi\Big(p_r P_{Rmax} - N(P_c+P_{DC})\Big)}}{\sqrt{\sum_{n=1}^{N}|\phi'_n|^2\Big(\sum_{u=1}^{U+1}|g_{n,u}|^2+\sigma_r^2\Big)}} < 1. \tag{41}$$

Proof. If $\boldsymbol{\Phi}'$ violates (29c), i.e., (39) does not hold with equality, which means that

$$\begin{aligned}&\mathbf{tr}\Big(\boldsymbol{\Phi}'\mathbf{H}^{\mathbf{B}}\boldsymbol{\Theta}_{\theta}^{\mathbf{DAC}}\mathbf{W}\mathbf{W}^{H}(\boldsymbol{\Phi}'\mathbf{H}^{\mathbf{B}})^{H} + \boldsymbol{\Phi}'\boldsymbol{\Phi}'^{H}\sigma_r^2\Big) >\\ &\xi\left(p_r P_{Rmax} - N(P_c + P_{DC})\right).\end{aligned} \tag{42}$$

From Proposition 3, $\boldsymbol{\Phi}$ satisfies (39). Thus, we have

$$\begin{aligned}&\mathbf{tr}\Big(\boldsymbol{\Phi}'\mathbf{H}^{\mathbf{B}}\boldsymbol{\Theta}_{\theta}^{\mathbf{DAC}}\mathbf{W}\mathbf{W}^{H}(\boldsymbol{\Phi}'\mathbf{H}^{\mathbf{B}})^{H} + \boldsymbol{\Phi}'\boldsymbol{\Phi}'^{H}\sigma_r^2\Big) >\\ &\mathbf{tr}\Big(\boldsymbol{\Phi}\mathbf{H}^{\mathbf{B}}\boldsymbol{\Theta}_{\theta}^{\mathbf{DAC}}\mathbf{W}\mathbf{W}^{H}(\boldsymbol{\Phi}\mathbf{H}^{\mathbf{B}})^{H} + \boldsymbol{\Phi}\boldsymbol{\Phi}^{H}\sigma_r^2\Big)\end{aligned} \tag{43}$$

By replacing (40) into (43), the following inequality is obtained

$$\frac{\xi\Big(p_r P_{Rmax} - N(P_c+P_{DC})\Big)}{\sum_{n=1}^{N}|\phi'_n|^2\Big(\sum_{u=1}^{U+1}|g_{n,u}|^2+\sigma_r^2\Big)} < 1, \tag{44}$$

which ensures that (41) is held. Therefore, the proof is completed. □

Then, from Lemma 1, we have $\boldsymbol{\Phi} = C\boldsymbol{\Phi}'$ with $C < 1$. This implies that $|\phi_n| = C|\phi'_n| < |\phi'_n| \le \psi_{max}$, i.e., $\boldsymbol{\Phi}$ satisfies (29d).

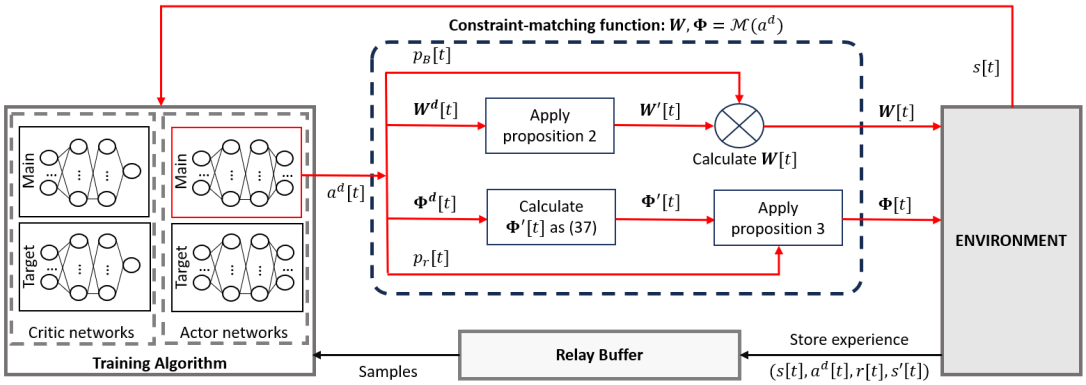


Fig. 2: Proposed DRL framework.

2) Proposed Deep Reinforcement Learning Algorithm: The proposed DRL framework is illustrated in Fig. 2. To decide action, the algorithm utilizes a NN named actor network, denoted by $\mu^{\theta_\mu}(s)$, where θ_μ is the NN's parameter. By denoting $a^d \triangleq \{p_B, p_r, \mathbf{W^d}, \mathbf{\Phi^d}\}$ is the action generalized from the actor network, at time slot t, it can be defined as

$$a^d[t] = \mu^{\theta_\mu}(s[t]). \tag{45}$$

Accordingly, the constraint-matching function is applied to ensure the action constraints, where the precoding and reflecting matrices are recalculated according to the constraint-matching function, expressed as

$$\mathbf{W}[t], \mathbf{\Phi}[t] = \mathcal{M}(a^d[t]). \tag{46}$$

The designed precoding and reflecting matrices are then used to interact with the environment. Then, a tuple of experience samples, including the state s, action a^d, reward r, and next state s', is stored into an experience buffer for the training process.

At each training step, a mini-batch of samples is randomly chosen from the buffer to update the NN parameters. The algorithm employs another NN called a critic network, denoted by $Q^{\theta_Q}(s,a)$, to train the actor network, where θ_Q is the NN parameter. The critic network expresses the action-value function of the corresponding state s and action a. Accordingly, the actor network parameter is updated using the policy gradient ascent function as

$$\nabla_{\theta_\mu} J = \frac{1}{B}\sum_{b=1}^{B}(\nabla_a Q^{\theta_Q}(s,a)|_{s=s_b, a=\mu^{\theta_\mu}(s_b)} \nabla_{\theta_\mu}\mu^{\theta_\mu}(s_b)), \tag{47}$$

where B is the training sample mini-batch size and s_b is the state of sample b. Then, the critic network parameter is updated by using the policy gradient descent for a loss function expressed as

$$L = \frac{1}{B}\sum_{b=1}^{B}(Q^{\theta_Q}(s_b, a_b) - y_b)^2, \tag{48}$$

where a_b is the action in sample b, $y_b = r_b + \gamma Q'^{\theta_{Q'}}(s'_b, \mu'^{\theta_{\mu'}}(s'_b))$ is the target value calculated by the reward r_b and the Q-value of the next state s'_b of sample b with a discount factor γ. In which, $\mu'^{\theta_{\mu'}}(s)$ and $Q'^{\theta_{Q'}}(s,a)$ are the target actor and critic networks, respectively, for enhancing the training stability [48]. The parameters of target networks are updated according to a soft-update strategy with a weight τ, expressed as

$$\begin{aligned} \theta_{\mu'} &\leftarrow \tau\theta_\mu + (1-\tau)\theta_{\mu'}, \\ \theta_{Q'} &\leftarrow \tau\theta_Q + (1-\tau)\theta_{Q'}. \end{aligned} \tag{49}$$

To explore the sample in the training process, the algorithm applies additional noise to the action. The noise follows the Ornstein-Uhlenbeck process [49], denoted by OU. Then, the action decided from the NN is expressed as

$$a^d[t] = \mu^{\theta_\mu}(s[t]) + OU[t]. \tag{50}$$

The proposed DRL-based algorithm is shown in Algorithm 1, which takes place in E episodes, each has T steps.

VI. Proposed Alternating Optimization Method

In this section, we propose an alternating optimization method to solve Problem (P1) named *Proposed-SBA*. Here, we decompose it into two sub-problems: *(i)* given the active reflecting matrix $\mathbf{\Phi}$, we optimize the precoding matrix $\mathbf{W}$; *(ii)* with the optimized precoding matrix $\mathbf{W}$, we solve the active reflecting matrix $\mathbf{\Phi}$. Using the above two-step optimization, the solution to (P1) can be obtained by iterative updating.

A. Precoding Matrix Optimization

With the given active reflecting matrix, the precoding matrix optimization problem can be formulated as

$$\text{(P1-A)}: \max_{\mathbf{W}} \quad EE \tag{51a}$$

$$\text{s.t.} \quad \text{tr}\left(\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{W}\mathbf{W}^H\right) \le P_{Bmax}, \tag{51b}$$

$$P_R \le P_{Rmax}. \tag{51c}$$

To deal with the nonconvex objective EE, we approximate (P1) by introducing auxiliary variables o, α^2, β, and ϱ that represent the EE metric, SE metric, BS power consumption,

and active RIS power consumption, respectively. Accordingly, problem (P1-A) is transformed to

$$\begin{aligned}
\text{(P2-A)}: \max_{\mathcal{V}_{2A}} \quad & o && (52a)\\
\text{s.t.} \quad & \frac{\alpha^2}{\beta+\varrho} \geq o, && (52b)\\
& r^c + \sum_{u\in\mathcal{U}} r_u^p \geq \alpha^2, && (52c)\\
& P_{AC} + \eta^{-1}\text{tr}\left(\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{W}\mathbf{W}^H\right) \leq \beta, && (52d)\\
& N(P_c + P_{DC}) \\
& \quad + \xi^{-1}\text{tr}\Big(\mathbf{\Phi}\mathbf{H}^{\mathbf{B}}\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{W}\mathbf{W}^H(\mathbf{\Phi}\mathbf{H}^{\mathbf{B}})^H \\
& \quad + \mathbf{\Phi}\mathbf{\Phi}^H\sigma_r^2\Big) \leq \varrho, && (52e)\\
& \text{tr}\left(\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{W}\mathbf{W}^H\right) \leq P_{Bmax}, && (52f)\\
& \varrho \leq P_{Rmax}, && (52g)
\end{aligned}$$

where $\mathcal{V}_{2A} \triangleq \{\mathbf{W}, o, \alpha, \beta, \varrho\}$, constraints (52b), (52c), (52d), and (52e) are obtained by transforming (P1-A) to (P2-A) with auxiliary variables, which hold with equality at optimum, (52f) and (52g) are the power constraints corresponding to (51b) and (51c). The challenge in this problem is caused by the nonconvexity of the constraints (52b) and (52c). To tackle this issue, we begin by applying the first-order Taylor approximation to approximate the left-hand side of constraint (52b), given as

$$\begin{aligned}
\frac{\alpha^2}{\beta+\varrho} &\geq \frac{(\alpha^{[i]})^2}{\beta^{[i]}+\varrho^{[i]}} + \frac{2\alpha^{[i]}}{\beta^{[i]}+\varrho^{[i]}}(\alpha-\alpha^{[i]}) \\
&\quad - \frac{(\alpha^{[i]})^2}{(\beta^{[i]}+\varrho^{[i]})^2}(\beta-\beta^{[i]}) - \frac{(\alpha^{[i]})^2}{(\beta^{[i]}+\varrho^{[i]})^2}(\varrho-\varrho^{[i]}) \\
&= \frac{2\alpha^{[i]}}{\beta^{[i]}+\varrho^{[i]}}\alpha - \frac{(\alpha^{[i]})^2}{(\beta^{[i]}+\varrho^{[i]})^2}(\beta+\varrho),
\end{aligned} \tag{53}$$

where $\alpha^{[i]}$, $\beta^{[i]}$, and $\varrho^{[i]}$ are the output values of α, β, and ϱ at i-th iteration, respectively. It can be verified that the function $\frac{\alpha^2}{\beta+\varrho}$ is convex for $\beta>0, \varrho>0$ since the corresponding Hessian matrix is positive semidefinite. Therefore, this function can be lower bounded by its first-order Taylor approximation. Accordingly, constraint (52b) can be rewritten by a convex constraint as

$$\frac{2\alpha^{[i]}}{\beta^{[i]}+\varrho^{[i]}}\alpha - \frac{(\alpha^{[i]})^2}{(\beta^{[i]}+\varrho^{[i]})^2}(\beta+\varrho) \geq o. \tag{54}$$

Next, we approximate constraint (52c) by introducing an auxiliary variable ρ to represent r^c. The constraint can be equivalently expressed as

$$\begin{aligned}
&\rho + \sum_{u\in\mathcal{U}} r_u^p \geq \alpha^2, && (55a)\\
&r_u^c \geq \rho, u\in\mathcal{U}, && (55b)
\end{aligned}$$

where (55b) is obtained by combining the approximated inequality, i.e., $r^c \geq \rho$, and (12).

For constraint (55a), we introduce variables $\delta_u, u\in\mathcal{U}$ that represent r_u^p, and express constraint (55a) as

$$\begin{aligned}
&\sum_{u\in\mathcal{U}} \log_2(\delta_u) \geq \alpha^2 - \rho, && (56a)\\
&1 + \frac{|\theta_u^{ADC}\mathbf{g}_u\mathbf{w}_u|^2}{\sum_{k\in\{\mathcal{U}\setminus u\}}|\theta_u^{ADC}\mathbf{g}_u\mathbf{w}_k|^2 + \text{NE}_u} \geq \delta_u, u\in\mathcal{U}, && (56b)
\end{aligned}$$

where $\mathbf{g}_u \triangleq \mathbf{h}_u^{\mathbf{R}}\mathbf{\Phi}\mathbf{H}^{\mathbf{B}}\mathbf{\Theta}_\theta^{\mathbf{DAC}}$. Then, by applying new variables $\varepsilon_u, u\in\mathcal{U}$, (56b) can be transformed to

$$\begin{aligned}
&|\mathbf{g}_u\mathbf{w}_u|^2 \geq (\delta_u-1)\varepsilon_u, u\in\mathcal{U}, && (57a)\\
&\varepsilon_u \geq \sum_{k\in\{\mathcal{U}\setminus u\}}|\mathbf{g}_u\mathbf{w}_k|^2 + \frac{\text{NE}_u}{(\theta_u^{ADC})^2}, u\in\mathcal{U}, && (57b)
\end{aligned}$$

where (57a) is remaining a nonconvex constraint. Based on [50, Eq. (B1)], the right-hand side of (57a) satisfies the following inequality

$$(\delta_u-1)\varepsilon_u \leq \frac{\varepsilon_u^{[i]}}{2(\delta_u^{[i]}-1)}(\delta_u-1)^2 + \frac{(\delta_u^{[i]}-1)}{2\varepsilon_u^{[i]}}\varepsilon_u^2, \tag{58}$$

where $\varepsilon_u^{[i]}$ and $\delta_u^{[i]}$ are the output values of ε_u and δ_u at i-th iteration, respectively. Besides, based on [51, Eq. (6a)], we approximate the left-hand side of (57a) as

$$|\mathbf{g}_u\mathbf{w}_u|^2 \geq 2\text{Re}\Big\{(\mathbf{g}_u\mathbf{w}_u^{[i]})^H\mathbf{g}_u\mathbf{w}_u\Big\} - |\mathbf{g}_u\mathbf{w}_u^{[i]}|^2. \tag{59}$$

Accordingly, (57a) can be rewritten by a convex constraint as

$$\begin{aligned}
&2\text{Re}\Big\{(\mathbf{g}_u\mathbf{w}_u^{[i]})^H\mathbf{g}_u\mathbf{w}_u\Big\} - |\mathbf{g}_u\mathbf{w}_u^{[i]}|^2 \\
&\quad \geq \frac{\varepsilon_u^{[i]}}{2(\delta_u^{[i]}-1)}(\delta_u-1)^2 + \frac{(\delta_u^{[i]}-1)}{2\varepsilon_u^{[i]}}\varepsilon_u^2, u\in\mathcal{U}.
\end{aligned} \tag{60}$$

From equations (56)–(60), constraint (55a) can be equivalently expressed by the following convex constraints

$$\begin{aligned}
&\sum_{u\in\mathcal{U}} \log_2(\delta_u) \geq \alpha^2 - \rho, && (61a)\\
&\varepsilon_u \geq \sum_{k\in\{\mathcal{U}\setminus u\}}|\mathbf{g}_u\mathbf{w}_k|^2 + \frac{\text{NE}_u}{(\theta_u^{ADC})^2}, u\in\mathcal{U}, && (61b)\\
&2\text{Re}\Big\{(\mathbf{g}_u\mathbf{w}_u^{[i]})^H\mathbf{g}_u\mathbf{w}_u\Big\} - |\mathbf{g}_u\mathbf{w}_u^{[i]}|^2 \\
&\quad \geq \frac{\varepsilon_u^{[i]}}{2(\delta_u^{[i]}-1)}(\delta_u-1)^2 + \frac{(\delta_u^{[i]}-1)}{2\varepsilon_u^{[i]}}\varepsilon_u^2, u\in\mathcal{U}. && (61c)
\end{aligned}$$

Similarly, following the same process with constraint (55a), constraint (55b) is equivalently expressed by

$$\begin{aligned}
&\log_2(\lambda_u) \geq \rho, u\in\mathcal{U}, && (62a)\\
&\nu_u \geq \sum_{k\in\mathcal{U}}|\mathbf{g}_u\mathbf{w}_k|^2 + \frac{\text{NE}_u}{(\theta_u^{ADC})^2}, u\in\mathcal{U}, && (62b)\\
&2\text{Re}\Big\{(\mathbf{g}_u\mathbf{w}_c^{[i]})^H\mathbf{g}_u\mathbf{w}_c\Big\} - |\mathbf{g}_u\mathbf{w}_c^{[i]}|^2 \\
&\quad \geq \frac{\nu_u^{[i]}}{2(\lambda_u^{[i]}-1)}(\lambda_u-1)^2 + \frac{(\lambda_u^{[i]}-1)}{2\nu_u^{[i]}}\nu_u^2, u\in\mathcal{U}. && (62c)
\end{aligned}$$

where $\lambda_u, \nu_u, u\in\mathcal{U}$ are the introduced auxilary variables, with $\lambda_u^{[i]}$ and $\nu_u^{[i]}$ being the corresponding values at i-th iteration. By

approximating (52b) and (52c) to convex constraints, problem (P2A) becomes a convex problem, expressed as

$$\max_{\mathcal{V}_{2Ao}} \quad o \tag{63a}$$
$$\text{s.t.} \quad (54), (61), (62), \tag{63b}$$
$$(52d), (52e), (52f), (52g), \tag{63c}$$

where $\mathcal{V}_{2Ao} \triangleq \{\mathbf{W}, o, \alpha, \beta, \varrho, \rho, \delta_u, \varepsilon_u, \lambda_u, \nu_u, u \in \mathcal{U}\}$. This convex problem can be solved using optimization tools like CVX, a package for specifying and solving convex programs [52], [53].

B. Active Reflecting Matrix Optimization

Given the precoding matrix $\mathbf{W}$, the active reflecting matrix optimization problem can be formulated from (P2) as

$$\text{(P1-B)}: \quad \max_{\mathbf{\Phi}} \quad EE \tag{64a}$$
$$\text{s.t.} \quad P_R \leq P_{Rmax}, \tag{64b}$$
$$|\phi_n| \leq \psi_{max}, n \in \mathcal{N}, \tag{64c}$$

which can be approximately transformed to

$$\text{(P2-B)}: \quad \max_{\mathcal{V}_{2B}} \quad x \tag{65a}$$
$$\text{s.t.} \quad \frac{\varsigma^2}{\varpi} \geq x, \tag{65b}$$
$$r_c + \sum_{u\in\mathcal{U}} r_u^p \geq \varsigma^2, \tag{65c}$$
$$P_{BS} + N(P_c + P_{DC}) + \xi^{-1}\text{tr}\Big(\mathbf{\Phi}\mathbf{H}^{\mathbf{B}}\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{W}\mathbf{W}^H(\mathbf{\Phi}\mathbf{H}^{\mathbf{B}})^H + \mathbf{\Phi}\mathbf{\Phi}^H\sigma_r^2\Big) \leq \varpi, \tag{65d}$$
$$\varpi \leq P_{BS} + P_{Rmax}, \tag{65e}$$
$$|\phi_n| \leq \psi_{max}, n \in \mathcal{N}, \tag{65f}$$

where $\mathcal{V}_{2B} \triangleq \{\mathbf{\Phi}, x, \varsigma, \varpi\}$, and x, ς, ϖ are the auxiliary variables. The challenge in solving this problem is the nonconvexity of constraints (65b) and (65c). To tackle this issue, we apply the same process as in sub-section VI-A. As a results, constraint (65b) can be rewritten by a convex constraint as

$$\frac{2\varsigma^{[i]}}{\varpi^{[i]}}\varsigma - \frac{(\varsigma^{[i]})^2}{(\varpi^{[i]})^2}\varpi \geq x, \tag{66}$$

Algorithm 2 *Proposed-SBA* algorithm.

1: **Initialize:**
2: $\mathbf{W}$, $\mathbf{\Phi}$;
3: **repeat**
4: Solve (63) $\rightarrow$ obtain $\mathbf{W}$.
5: Solve (68) $\rightarrow$ obtain $\mathbf{\Phi}$.
6: Update $\mathbf{W}$, $\mathbf{\Phi}$.
7: **until** convergence
8: **return** $\Phi^*, \mathbf{W}^*$

and constraint (65c) can be approximately expressed as the following constraints

$$\sum_{u\in\mathcal{U}} \log_2(\Lambda_u) \geq \varsigma^2 - \wp, \tag{67a}$$
$$\varkappa_u \geq \sum_{k\in\{\mathcal{U}\setminus u\}} |\mathbf{g}_u\mathbf{w}_k|^2 + \frac{\text{NE}_u}{(\theta_u^{ADC})^2}, u \in \mathcal{U}, \tag{67b}$$
$$2\text{Re}\Big\{(\mathbf{g}_u^{[i]}\mathbf{w}_u)^H\mathbf{g}_u\mathbf{w}_u\Big\} - |\mathbf{g}_u^{[i]}\mathbf{w}_u|^2 \geq \frac{\varkappa_u^{[i]}}{2(\Lambda_u^{[i]}-1)}(\Lambda_u-1)^2 + \frac{(\Lambda_u^{[i]}-1)}{2\varkappa_u^{[i]}}\varkappa_u^2, u \in \mathcal{U}, \tag{67c}$$
$$\log_2(\epsilon_u) \geq \wp, u \in \mathcal{U}, \tag{67d}$$
$$\ell_u \geq \sum_{k\in\mathcal{U}} |\mathbf{g}_u\mathbf{w}_k|^2 + \frac{\text{NE}_u}{(\theta_u^{ADC})^2}, u \in \mathcal{U}, \tag{67e}$$
$$2\text{Re}\Big\{(\mathbf{g}_u^{[i]}\mathbf{w}_c)^H\mathbf{g}_u\mathbf{w}_c\Big\} - |\mathbf{g}_u^{[i]}\mathbf{w}_c|^2 \geq \frac{\ell_u^{[i]}}{2(\epsilon_u^{[i]}-1)}(\epsilon_u-1)^2 + \frac{(\epsilon_u^{[i]}-1)}{2\ell_u^{[i]}}\ell_u^2, u \in \mathcal{U}, \tag{67f}$$

where $\mathbf{g}_u^{[i]} \triangleq \mathbf{h}_u^{\mathbf{R}}\mathbf{\Phi}^{[i]}\mathbf{H}^{\mathbf{B}}\mathbf{\Theta}_\theta^{\mathbf{DAC}}$, and $\wp, \Lambda_u, \varkappa_u, \epsilon_u, \ell_u, u \in \mathcal{U}$ are the introduced auxiliary variables. Accordingly, (P2-B) becomes a convex problem with new convex constraints, expressed as

$$\max_{\mathcal{V}_{2Bo}} \quad x \tag{68a}$$
$$\text{s.t.} \quad (66), (67), \tag{68b}$$
$$(65d), (65e), (65f), \tag{68c}$$

where $\mathcal{V}_{2Bo} \triangleq \{\mathbf{\Phi}, x, \varsigma, \varpi, \wp, \Lambda_u, \varkappa_u, \epsilon_u, \ell_u, u \in \mathcal{U}\}$. This convex problem can be solved using convex optimization tools. The proposed-SBA algorithm is summarized in Algorithm 2.

C. Convergence Analysis

In this subsection, convergence analysis for Algorithm 2 is provided based on the following proposition.

Proposition 4. Algorithm 2 is guaranteed to converge to a sub-optimal solution to problem (P1).

Proof. Let $f(\mathbf{W}, \mathbf{\Phi}), f_{\mathbf{\Phi}}(\mathcal{V}_{2Ao})$, and $f_{\mathbf{W}}(\mathcal{V}_{2Bo})$ denote the objective function of problem (P1), problem (63) for given $\mathbf{\Phi}$, and problem (68) for given $\mathbf{W}$, respectively. At the i-th iteration of Algorithm 2, for given $\mathbf{\Phi}^{[i]}$, we have

$$f(\mathbf{W}^{[i]}, \mathbf{\Phi}^{[i]}) \overset{(f)}{=} f_{\mathbf{\Phi}^{[i]}}(\mathcal{V}_{2Ao}^{[i]}) \overset{(g)}{\leq} f_{\mathbf{\Phi}^{[i]}}(\mathcal{V}_{2Ao}^{[i+1]}) = f(\mathbf{W}^{[i+1]}, \mathbf{\Phi}^{[i]}), \tag{69}$$

where (f) is due to the fact that the first-order Taylor approximation in (54) and the approximations in (61c), (62c) are tight at the current point $\mathcal{V}_{2Ao}^{[i]}$; (g) is due to the fact that $\mathcal{V}_{2Ao}^{[i+1]}$ is

TABLE IV: Environmental parameters

Parameter	Value
b^{DAC}, b^{ADC}	2–10 bits
Active RIS size	$N = 16(4 \times 4)$
P_{Bmax}	0-24 dBm
σ_r^2, σ_u^2	-174 dBm/Hz
Bandwidth	10 MHz
P_C	-10 dBm
P_{DC}	-5 dBm
ξ	0.8
P_{Rmax}	20 dBm
ψ_{max}	10
Rician factors	1000

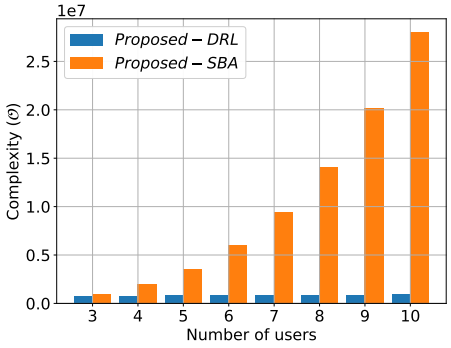


Fig. 3: Complexity according to the number of users

optimal at i-th iteration. For given $\mathbf{W}^{[i+1]}$, performing the same argument yields

$$f(\mathbf{W}^{[i+1]}, \mathbf{\Phi}^{[i]}) = f_{\mathbf{W}^{[i+1]}}(\mathcal{V}_{2Bo}^{[i]}) \leq f_{\mathbf{W}^{[i+1]}}(\mathcal{V}_{2Bo}^{[i+1]}) = f(\mathbf{W}^{[i+1]}, \mathbf{\Phi}^{[i+1]}). \quad (70)$$

From (69) and (70), we have

$$f(\mathbf{W}^{[i]}, \mathbf{\Phi}^{[i]}) \leq f(\mathbf{W}^{[i+1]}, \mathbf{\Phi}^{[i+1]}), \quad (71)$$

which means that the sequence of the objective values of problem (P1) is monotonically non-decreasing. In addition, the system EE is upper bounded. Thus, Algorithm 2 is guaranteed to converge to a sub-optimal solution to problem (P1). □

VII. Numerical Results

To evaluate the proposed solutions and system, we simulate an environment consisting of a BS equipped with 8 antennas serving 10 users randomly distributed in a range of 50-to-200 meters from the BS. Considering BS as the coordinate origin, the active RIS is deployed at a position of (10, 20, 10) meters. The channels between the BS and RIS and between RIS and users are modeled following the Rician fading channel model, represented as

$$\mathbf{H}^{\mathbf{B}} = L_B \left(\sqrt{\frac{f_1}{f_1+1}} \mathbf{H}^{\mathbf{B}^{LoS}} + \sqrt{\frac{1}{f_1+1}} \mathbf{H}^{\mathbf{B}^{NLoS}} \right),$$
$$\mathbf{h}_u^{\mathbf{R}} = L_{R,u} \left(\sqrt{\frac{f_2}{f_2+1}} \mathbf{h}_u^{\mathbf{R}^{LoS}} + \sqrt{\frac{1}{f_2+1}} \mathbf{h}_u^{\mathbf{R}^{NLoS}} \right), \quad (72)$$

where f_1 and f_2 are the Rician factors of the BS-RIS and RIS-user channels, respectively; $\mathbf{H}^{\mathbf{B}^{LoS}}$ and $\mathbf{h}_u^{\mathbf{R}^{LoS}}$ are the line-of-sight (LoS) components of the channels, which are modeled according to array response vectors described in [54, sub-section III-B]; $\mathbf{H}^{\mathbf{B}^{NLoS}}$ and $\mathbf{h}_u^{\mathbf{R}^{NLoS}}$ are the corresponding non-line-of-sight (NLoS) components modeled as Rayleigh fading variables; L_B and $L_{R,u}$ are the corresponding path-loss, calculated as: $L = -L_0 - 10\upsilon \log_{10}\left(\frac{d}{d_0}\right)$ [32], with $L_0 = 30$ dB, $\upsilon = 2.2$, $d_0 = 1$ m, and d is the distance link. Table IV summarizes the other environmental parameters. NNs in the DRL algorithm have two hidden layers, with 1024 and 512 nodes in the first and second hidden layers, respectively. The experience buffer has 10^5 entries, and we set $\gamma = 0.99$, $\tau = 0.01$, and $B = 16$.

A. Complexity Analysis

In this sub-section, we analyze the complexity of the proposed approaches and evaluate their feasibility in practical scenarios.

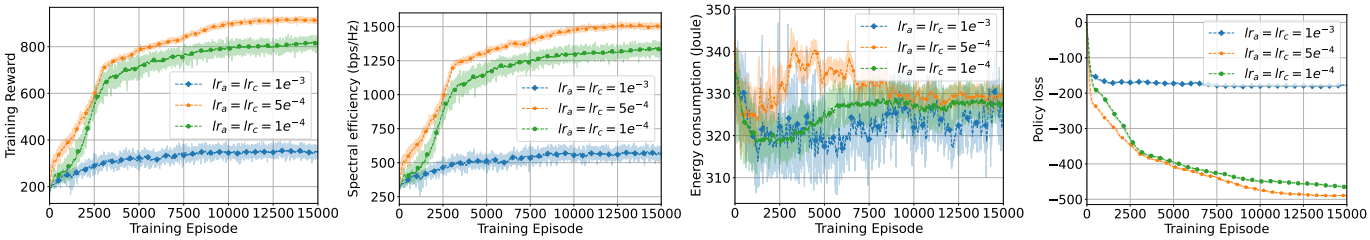
1) Complexity of the Proposed Alternating Algorithm: The main complexity of Algorithm 2 lies in solving the convex optimization problems (63) and (68). The number of variables in (63) is $(U+1)\times(M+4)$. Similarly, the number of variables in (68) is $4U+N+3$. At each iteration step, each sub-problem can be solved using a CVX toolbox. According to [55], [56], the computational complexity of solving (63) and (68) are $\mathcal{O}\Big(((U+1)\times(M+4))^{3.5}\Big)$ and $\mathcal{O}\Big((4U+N+3)^{3.5}\Big)$, respectively. Consequently, the overall complexity of Algorithm 2 in each iteration is $\mathcal{O}\Big(\big((U+1)\times(M+4)\big)^{3.5}+(4U+N+3)^{3.5}\Big)$ $\approx \mathcal{O}\Big((UM)^{3.5}+N^{3.5}\Big)$.

2) Complexity of the Proposed DRL Algorithm: Considering that the DRL approach's training process is conducted in advance, we calculate the computational complexity of the proposed DRL framework when interacting with the environment based on the actor network. According to [42], [44], [54], the complexity can be calculated as $\mathcal{O}\Big(\sum_{l=1}^{L}\zeta_{l-1}\zeta_l\Big)$. In which, ζ_l is the number of nodes at layer l of the network, where the number of nodes in the input layer ($l = 0$) and output layer ($l = L$) are equal to the number of entries in the state space $s[t]$ and action space $a^d[t]$, respectively, expressed as

$$\zeta_0 = N \times (U+M),$$
$$\zeta_L = M \times (U+1) + N + 2. \quad (73)$$

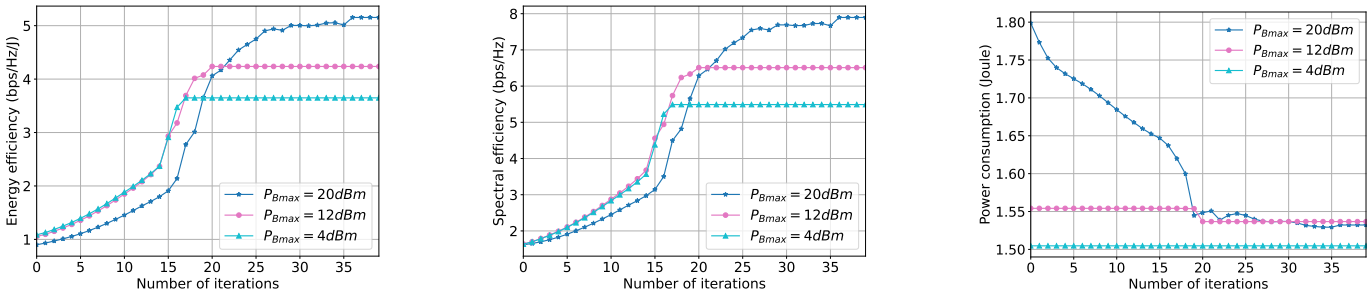
Accordingly, the computational complexity of the proposed DRL algorithm is $\mathcal{O}(NU + NM + MU)$.

3) Complexity Comparison: Based on the above analyses, the proposed approaches have polynomial complexities, which are scalable to the environment. For instance, the complexity of the alternating-based approach are $\mathcal{O}(U^{3.5})$, $\mathcal{O}(M^{3.5})$, and $\mathcal{O}(N^{3.5})$ in accordance with the number of users, BS antennas, and reflecting elements, while the DRL-based approach has the linear complexity with these parameters. As illustrated in Fig. 3, the complexity of the *Proposed-SBA* algorithm dramatically increases with the number of users, whereas the *Proposed-DRL* remains relatively minor. Therefore, the *Proposed-SBA* algorithm is mainly suitable for small-scale environments, but even in these cases, it requires significant computational resources, making high-performance computing



(a) Training reward. (b) Spectral efficiency. (c) Power consumption. (d) Policy loss.

Fig. 4: Training results of the *Proposed-DRL* algorithm with different learning rate.



(a) Energy efficiency. (b) Spectral efficiency. (c) Power consumption.

Fig. 5: Convergence of the *Proposed-SBA* algorithm with different transmit power.

systems necessary for efficient execution. On the other hand, the *Proposed-DRL* algorithm, with its lower computational complexity, is better suited for practical applications. However, its reliance on fixed feedforward neural networks limits its adaptability to changes in state and action scales, requiring time-consuming retraining when these scales change. Thus, the DRL approach is more appropriate for environments with stable scales than dynamic ones. Future research could explore integrating alternative neural network architectures, such as liquid neural networks, to enhance adaptability and effectiveness in more dynamic settings.

B. Convergence Analysis

First, we evaluate the convergence of the proposed DRL framework (*Proposed-DRL*) in different learning rates, where the actor learning rate (lr_a) and critic learning rate (lr_c) are experimented with three values of $1e^{-3}$, $5e^{-4}$, and $1e^{-4}$. After training the agent in 15000 episodes, each with 200 steps, the case $lr_a = lr_c = 5e^{-4}$ results in the best reward, where it hits the convergence value of about 900 after 12000 episodes, as illustrated in Fig. 4a. The case $lr_a = lr_c = 1e^{-4}$ has a reward value of about 800 after 15000 episodes because, with an undersized learning rate, the parameter updates with a tiny value after each step, lengthening the training time. Otherwise, in the remaining case with a high learning rate, the high value of each updating step makes the model fluctuate but cannot reach the optimal value. The change in reward is in accordance with the increase in SE and the stability of power consumption, as presented in Figs. 4b and 4c. Moreover, we evaluate the policy loss value during the training period to observe the training process of the actor network. Based on (47), we measure the policy loss, L^{μ}, by

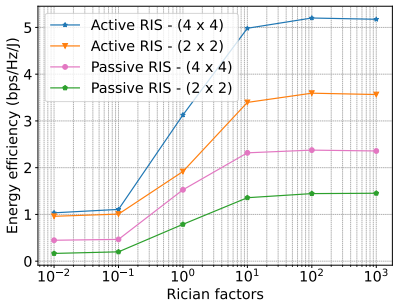
$$L^{\mu} = -\frac{1}{B}\sum_{b=1}^{B}(Q^{\theta_Q}(s_b, \mu^{\theta_\mu}(s_b))). \tag{74}$$

As shown in Fig. 4d, the policy loss gradually reduces in the training process, demonstrating an increase in the Q-value of action over time, i.e., the model is trained well. Also, the result is consistent with the training reward, where the case $lr_a = lr_c = 5e^{-4}$ results the best value and converges after about 12000 episodes. Hence, we use the model with the best performance for the evaluation.

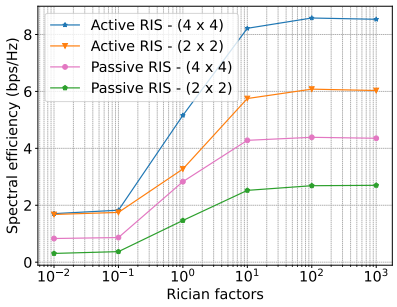
Second, we assess the proposed SCA-based alternating optimization method (*Proposed-SBA*) convergence in different environmental scenarios, i.e., maximum BS transmit power. Here, we set the convergence threshold to 99% of the previous value, and the initialize values are set to satisfy all the problem constraints. As illustrated in Fig. 5, all three cases converge after 40 iterations, with the case of the highest power providing the best result. Additionally, the convergence of the EE is consistent with the increase in the SE and the decrease in the power consumption.

C. Performance Evaluation

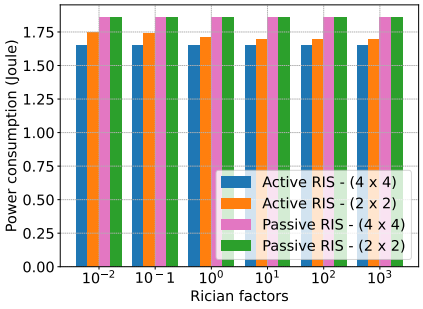
This sub-section begins with an analysis of the proposed system performance in various fading environments by varying the Rician factors. Besides, we assess the efficiency of the active RIS system by comparing it with the passive RIS at different RIS sizes. As observed in Fig. 6, the EE increases according to the Rician factor values. The small Rician factors



(a) Energy efficiency.



(b) Spectral efficiency.



(c) Power consumption.

Fig. 6: RIS system in fading environment.

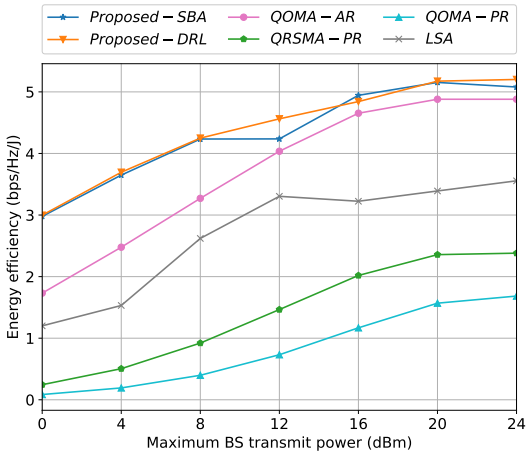


Fig. 7: Energy efficiency with different BS transmit power

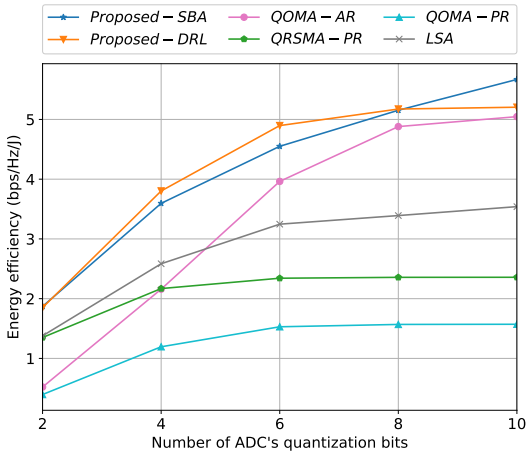


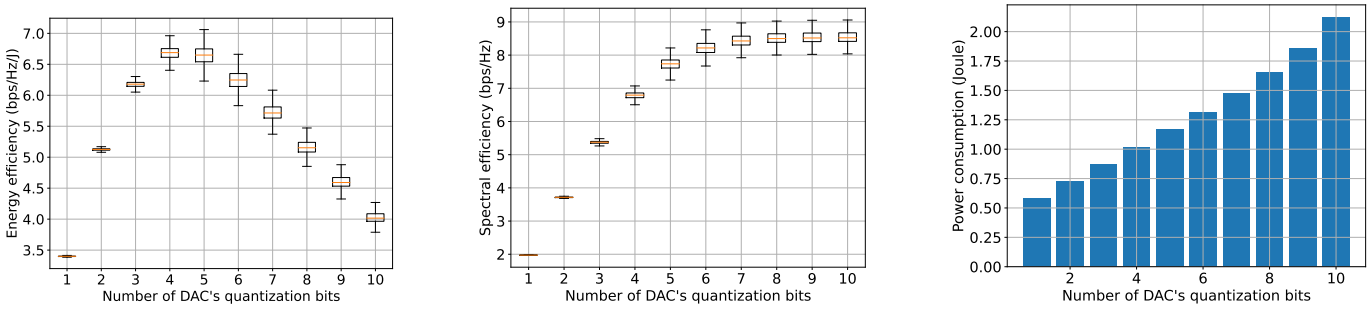
Fig. 8: Energy efficiency with different ADC's resolution

lead to a high effect of the NLoS components in the communication channels, boosting the chaos in the signal and thus reducing the system performance. The results reveal that the system works well when the Rician factors are higher than 10, reducing performance by about 80%, 7.6%, and 39.5% when the Rician factors are 10^{-2}, 10^{-1}, and 10^{0}, respectively. Also, the results demonstrate the superiority of the active RIS schemes over passive RIS schemes, with the EE values more than 2.2 times higher. Interestingly, Fig. 6c indicates that the system's power consumption ($P_R + P_B$) with passive RIS is higher than with active RIS. This is because active RIS enhances signal propagation more effectively, significantly reducing the BS's transmission power. Although an active RIS can consume more power than a passive RIS, reducing BS transmission power can offset this, resulting in lower overall system power consumption with active RIS. Additionally, large-scale RIS can reduce power consumption compared to small-scale RIS by providing more effective signal manipulation, allowing for even lower transmission power from the BS and reducing overall system power consumption. In addition, using the larger size of the RIS (4×4) can improve the performance, where the EE rises by over 40% compared to the smaller size (2×2).

Then, we simulate the following benchmark schemes to assess the effectiveness of the proposed frameworks

- **Quantization with orthogonal multiple access (OMA) and Active RIS (*QOMA-AR*)**: We conduct this scheme to compare the performance of RSMA to the OMA schemes in the quantized system. According to [57], we simulate the OMA system using the orthogonal frequency-division multiple-access (OFDMA) technique, where the BS allocates a dedicated part of the bandwidth for each user with no inter-user interference.
- **Quantization with RSMA and Passive RIS (*QRSMA-PR*)**: This scheme uses the combination of RSMA with the passive RIS, as considered in [30], [31], in the quantized system.
- **Quantization with OMA and passive RIS (*QOMA-PR*)**: Similar to the *QOMA-AR* scheme, this system applies OFDMA in the transmission but combines it with the passive RIS.
- **Optimization using local search-based algorithm (*LSA*)**: We simulate this scheme as a sub-optimal solution to validate the proposed algorithms by discretizing the action values and selecting the best strategy at each time slot. Given the high action space dimensionality, we apply a low-complexity search method from [58, Algorithm 2] to solve the problem efficiently.

Fig. 7 illustrates the EE for all the schemes with different maximum transmit power, where we increase the P_{Bmax} from 0 to 24 dBm. As a result, the increment of the transmit power yields a rise in the EE because of the increase in SE, and the



(a) Energy efficiency. (b) Spectral efficiency. (c) Power consumption.

Fig. 9: System performance with different DAC's resolution.

EE reaches the peak value at $P_{Bmax} = 20$ dBm. The results show that the proposed active RIS and RSMA combination scheme outperforms other benchmarks, where it is about 25.98%, 2.26 times, and 3.51 times compared to *QOMA-AR*, *QRSMA-PR*, and *QOMA-PR* schemes, respectively. The improvement of active RIS is also demonstrated in this result, which is illustrated by the superiority of the proposed scheme and *QOMA-AR* compared to the corresponding *QRSMA-PR* and *QOMA-PR* schemes. Besides, the proposed DRL and alternating algorithms obtain approximately the same performance and demonstrate their superiority, at about 49.7% better than the *LSA* scheme. Then, we estimate the EE under various quantization levels at the receivers, where we vary the number of ADCs quantization bits from 2 to 10. As depicted in Fig. 8, the EE increases in accordance with the quantization level because using high-resolution quantizers lowers the quantization noise, reducing the signal distortion and thus improving transmission efficiency. The system performs well when the quantization levels are higher than 8 bits. In particular, the EE rises about three times when increasing the quantization level from 2 to 8 bits. Besides, the OMA schemes (*QOMA-AR* and *QOMA-PR*) perform poorly in low-resolution cases (2 and 4 bits), implying that RSMA is better than OMA in enhancing systems with high quantization noise.

Finally, we analyze the effect of the quantization level at the transmitter by varying the number of DACs quantization bits from 1 to 10. The results are measured in 3000 consecutive steps. As displayed in Fig. 9, the EE initially rises as the number of bits increases from 1 to 4 but subsequently decreases. The rationale behind this observation is depicted in Figs. 9b and 9c, where the SE increases as the quantization level rises and stabilizes when the resolution surpasses 7 bits. However, the power consumption consistently increases with the quantization level. Hence, the stability of the SE and the increase in power consumption cause a deduction in the EE. Besides, Fig. 9a reveals that the case of 4-bit resolution at the DAC gives the best EE. However, the SE in that case is only about 6.78 bps/Hz and is not the best SE case. Therefore, in practical scenarios, a trade-off between the EE, SE, and power consumption should be carefully considered in accordance with the system requirements.

VIII. Conclusion

This study concentrated on improving the performance of a quantized downlink multi-user MISO system, where the transceivers are equipped with low-resolution converters, by combining the effectiveness of active RIS and RSMA in the transmission. In particular, we aimed to maximize the energy efficiency in this quantized system, which is calculated based on the spectral efficiency and system power consumption. To do so, we formulated an energy efficiency maximization problem considering the constraints of the BS precoding and active RIS reflecting matrices. The nonconvexity of the objective presented a tough challenge for solving this problem. To overcome this issue, we proposed solutions based on reinforcement learning and alternating optimization. The first approach presented the problem as a reinforcement learning-based problem solvable by applying a DRL algorithm named DDPG. However, the original DDPG algorithm may not satisfy the problem constraints and thus distort the action accuracy. Therefore, we proposed a constraint-matching function to be integrated into the DDPG algorithm, resulting in a DRL framework that efficiently designs the action while ensuring the problem constraints. The second approach divided the problem into two sub-problems: precoding matrix optimization and active reflecting matrix optimization. Each was solved using the SCA-based method, involving several mathematical steps to approximate it to a convex problem that is resolvable using convex optimization tools. The simulation results demonstrated the convergence of the proposed solutions. In addition, we proved the effectiveness of the proposed system and algorithms compared to other state-of-the-art benchmark schemes under various scenarios. Finally, we analyzed the influence of the quantization levels on the system performance. The results also show the importance of the trade-off between EE, SE, and power consumption in practical applications.

In multi-antenna systems, SDMA (space-division multiple access) can provide an efficient approach by using spatial separation to allow multiple users to share the same frequency band without interference. Additionally, RIS deployment can enhance channel orthogonality, potentially narrowing the performance gap between SDMA and RSMA. However, this advantage depends on the environment. In dense settings with close or even co-located users, RSMA remains essential for

interference mitigation. In contrast, SDMA combined with RIS can perform effectively in distributed environments. This trade-off merits further research to explore how RIS characteristics affect the relative performance of RSMA and SDMA.

APPENDIX A
POWER CONSTRAINT REFORMULATION

The expectation of the product $\mathbf{x}^q(\mathbf{x}^q)^H$ is expressed as

$$\mathbb{E}\big[\mathbf{x}^q(\mathbf{x}^q)^H\big] = \mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{W}\mathbf{W}^H(\mathbf{\Theta}_\theta^{\mathbf{DAC}})^H + \mathbf{R}^{\mathbf{DAC}}. \quad \text{(A.1)}$$

Then, taking the trace operator on both sides of (A.1) yields

$$\mathrm{tr}(\mathbb{E}\big[\mathbf{x}^q(\mathbf{x}^q)^H\big]) = \mathrm{tr}\Big(\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{W}\mathbf{W}^H(\mathbf{\Theta}_\theta^{\mathbf{DAC}})^H + \mathbf{R}^{\mathbf{DAC}}\Big), \quad \text{(A.2)}$$

where $\mathbf{R}^{\mathbf{DAC}}$ is calculated as

$$\begin{aligned}\mathbf{R}^{\mathbf{DAC}} &= \mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{\Theta}_\vartheta^{\mathbf{DAC}}\mathrm{diag}\left(\mathbb{E}\big[\mathbf{x}\mathbf{x}^H\big]\right)\\ &= \mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{\Theta}_\vartheta^{\mathbf{DAC}}\mathrm{diag}\Big(\mathbf{W}\mathbf{W}^H\Big)\end{aligned} \quad \text{(A.3)}$$

Using the relation $\theta_m^{DAC} = 1 - \vartheta_m^{DAC}$ yields $\mathbf{\Theta}_\vartheta^{\mathbf{DAC}} = \mathbf{I}_M - \mathbf{\Theta}_\theta^{\mathbf{DAC}}$. Then, (A.3) can be rewritten as

$$\mathbf{R}^{\mathbf{DAC}} = \mathbf{\Theta}_\theta^{\mathbf{DAC}}(\mathbf{I}_M - \mathbf{\Theta}_\theta^{\mathbf{DAC}})\mathrm{diag}\Big(\mathbf{W}\mathbf{W}^H\Big). \quad \text{(A.4)}$$

Substituting (A.4) into (A.2) yields

$$\begin{aligned}\mathrm{tr}\left(\mathbb{E}\left[\mathbf{x}^q(\mathbf{x}^q)^H\right]\right) = \mathrm{tr}&\Big(\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{W}\mathbf{W}^H\Big(\mathbf{\Theta}_\theta^{\mathbf{DAC}}\Big)^H\Big)\\ &+ \mathrm{tr}\Big(\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathrm{diag}\Big(\mathbf{W}\mathbf{W}^H\Big)\Big)\\ &- \mathrm{tr}\Big(\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathrm{diag}\Big(\mathbf{W}\mathbf{W}^H\Big)\Big).\end{aligned} \quad \text{(A.5)}$$

For a given real diagonal matrix, $\mathbf{\Theta}_\theta^{DAC} \triangleq \mathrm{diag}(\theta_1^{DAC}, \theta_2^{DAC}, \ldots, \theta_M^{DAC})$, the first term in (A.5) is transformed into

$$\mathrm{tr}\Big(\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{W}\mathbf{W}^H\Big(\mathbf{\Theta}_\theta^{\mathbf{DAC}}\Big)^H\Big) = \mathrm{tr}\Big(\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{W}\mathbf{W}^H\Big), \quad \text{(A.6)}$$

and the last term in (A.5) is transformed into

$$\mathrm{tr}\Big(\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathrm{diag}\Big(\mathbf{W}\mathbf{W}^H\Big)\Big) = \mathrm{tr}\Big(\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{W}\mathbf{W}^H\Big). \quad \text{(A.7)}$$

From (A.6) and (A.7), we observe that these two terms are equal. Thus, in (A.5), they are canceled out and we remain

$$\begin{aligned}\mathrm{tr}\left(\mathbb{E}\left[\mathbf{x}^q(\mathbf{x}^q)^H\right]\right) &= \mathrm{tr}\Big(\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathrm{diag}\Big(\mathbf{W}\mathbf{W}^H\Big)\Big)\\ &= \mathrm{tr}\Big(\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{W}\mathbf{W}^H\Big).\end{aligned} \quad \text{(A.8)}$$

As a result, the power constraint (5) is rewritten as

$$\mathrm{tr}\Big(\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{W}\mathbf{W}^H\Big) \le P_{Bmax}. \quad \text{(A.9)}$$

This completes the proof.

APPENDIX B
PROOF OF PROPOSITION 3

In this part, we prove that the normalized active reflecting matrix, $\mathbf{\Phi} \triangleq \mathrm{diag}\,(\phi_1, \phi_2, \ldots, \phi_N)$, satisfies constraint (29c), that is, the equation (39) is held. By letting $\mathbf{G} \triangleq \mathbf{H}^{\mathbf{B}}\mathbf{\Theta}_\theta^{\mathbf{DAC}\frac{1}{2}}\mathbf{W} \in \mathbb{C}^{N\times(U+1)}$, we first reform the left side of (39) as

$$\begin{aligned}&\mathbf{tr}\Big(\mathbf{\Phi}\mathbf{H}^{\mathbf{B}}\mathbf{\Theta}_\theta^{\mathbf{DAC}}\mathbf{W}\mathbf{W}^H(\mathbf{\Phi}\mathbf{H}^{\mathbf{B}})^H + \mathbf{\Phi}\mathbf{\Phi}^H\sigma_r^2\Big)\\ &= \mathbf{tr}\left(\mathbf{\Phi}\mathbf{G}\mathbf{G}^H\mathbf{\Phi}^H\right) + \sigma_r^2\mathbf{tr}\left(\mathbf{\Phi}\mathbf{\Phi}^H\right)\\ &= \| \mathbf{\Phi}\mathbf{G} \|_2^2 + \sigma_r^2 \| \mathbf{\Phi} \|_2^2\\ &= \sum_{n=1}^{N}\sum_{u=1}^{U+1}|\phi_n g_{n,u}|^2 + \sigma_r^2\sum_{n=1}^{N}|\phi_n|^2\\ &= \sum_{n=1}^{N}|\phi_n|^2\left(\sum_{u=1}^{U+1}|g_{n,u}|^2 + \sigma_r^2\right).\end{aligned} \quad \text{(B.2)}$$

We replace (40) to (B.2) using (B.1). Then, (39) is obtained, which proves the proposition.

REFERENCES

[1] J. Choi, G. Lee, A. Alkhateeb, A. Gatherer, N. Al-Dhahir, and B. L. Evans, "Advanced receiver architectures for millimeter-wave communications with low-resolution ADCs," *IEEE Communications Magazine*, vol. 58, no. 8, pp. 42–48, 2020.

[2] J. Choi, J. Park, and N. Lee, "Energy efficiency maximization precoding for quantized massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 21, no. 9, pp. 6803–6817, 2022.

[3] O. De Candido, H. Jedda, A. Mezghani, A. L. Swindlehurst, and J. A. Nossek, "Reconsidering linear transmit signal processing in 1-bit quantized multi-user MISO systems," *IEEE Transactions on Wireless Communications*, vol. 18, no. 1, pp. 254–267, 2019.

[4] S. Jacobsson, G. Durisi, M. Coldrey, and C. Studer, "Linear precoding with low-resolution DACs for massive MU-MIMO-OFDM downlink," *IEEE Transactions on Wireless Communications*, vol. 18, no. 3, pp. 1595–1609, 2019.

[5] Y. Guan, J. Deng, and X. Cheng, "On the accuracy of wideband localization in quantized uplink massive MIMO systems," *IEEE Communications Letters*, vol. 26, no. 7, pp. 1558–1562, 2022.

[6] S. Gong, X. Lu, D. T. Hoang, D. Niyato, L. Shu, D. I. Kim, and Y.-C. Liang, "Toward smart wireless communications via intelligent reflecting surfaces: A contemporary survey," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 4, pp. 2283–2314, 2020.

[7] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1546–1577, 2021.

[8] Z. Zhang, L. Dai, X. Chen, C. Liu, F. Yang, R. Schober, and H. V. Poor, "Active RIS vs. passive RIS: Which will prevail in 6G?" *IEEE Transactions on Communications*, vol. 71, no. 3, pp. 1707–1725, 2023.

[9] K. Zhi, C. Pan, H. Ren, K. K. Chai, and M. Elkashlan, "Active RIS versus passive RIS: Which is superior with the same power budget?" *IEEE Communications Letters*, vol. 26, no. 5, pp. 1150–1154, 2022.

[10] W. Lv, J. Bai, Q. Yan, and H. M. Wang, "RIS-assisted green secure communications: Active RIS or passive RIS?" *IEEE Wireless Communications Letters*, vol. 12, no. 2, pp. 237–241, 2023.

[11] Y. Chen, Y. Wang, Z. Wang, and P. Zhang, "Robust beamforming for active reconfigurable intelligent omni-surface in vehicular communications," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 10, pp. 3086–3103, 2022.

[12] B. Lyu, C. Zhou, S. Gong, D. T. Hoang, and Y.-C. Liang, "Robust secure transmission for active RIS enabled symbiotic radio multicast communications," *IEEE Transactions on Wireless Communications*, vol. 22, no. 12, pp. 8766–8780, 2023.

[13] Z. Yu, H. Ren, C. Pan, G. Zhou, B. Wang, M. Dong, and J. Wang, "Active RIS aided ISAC systems: Beamforming design and performance analysis," *IEEE Transactions on Communications*, pp. 1–1, 2023.

$$\mathbf{tr}\Big(\mathbf{\Phi H^B \Theta_\theta^{DAC} W W}^H(\mathbf{\Phi H^B})^H + \mathbf{\Phi \Phi}^H\sigma_r^2\Big) = \sum_{n=1}^{N}\left|\frac{\phi_n'\sqrt{\xi\Big(p_r P_{Rmax} - N(P_c+P_{DC})\Big)}}{\sqrt{\sum_{n=1}^{N}|\phi_n'|^2\Big(\sum_{u=1}^{U+1}|gn,u|^2+\sigma_r^2\Big)}}\right|^2\left(\sum_{u=1}^{U+1}|g_{n,u}|^2+\sigma_r^2\right)$$
$$= \frac{\sum_{n=1}^{N}\left|\phi_n'\sqrt{\xi\Big(p_r P_{Rmax} - N(P_c+P_{DC})\Big)}\right|^2}{\sum_{n=1}^{N}|\phi_n'|^2\Big(\sum_{u=1}^{U+1}|g_{n,u}|^2+\sigma_r^2\Big)}\left(\sum_{u=1}^{U+1}|g_{n,u}|^2+\sigma_r^2\right) = \xi\Big(p_r P_{Rmax} - N(P_c+P_{DC})\Big). \tag{B.1}$$

[14] R. K. Fotock, A. Zappone, and M. D. Renzo, "Energy efficiency optimization in RIS-aided wireless networks: Active versus nearly-passive RIS with global reflection constraints," *IEEE Transactions on Communications*, vol. 72, no. 1, pp. 257–272, 2024.

[15] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 49–60, 1981.

[16] B. Rimoldi and R. Urbanke, "A rate-splitting approach to the gaussian multiple-access channel," *IEEE Transactions on Information Theory*, vol. 42, no. 2, pp. 364–375, 1996.

[17] Y. Mao, O. Dizdar, B. Clerckx, R. Schober, P. Popovski, and H. V. Poor, "Rate-splitting multiple access: Fundamentals, survey, and future research trends," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 4, pp. 2073–2126, 2022.

[18] B. Clerckx, Y. Mao, E. A. Jorswieck, J. Yuan, D. J. Love, E. Erkip, and D. Niyato, "A primer on rate-splitting multiple access: Tutorial, myths, and frequently asked questions," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 5, pp. 1265–1308, 2023.

[19] Z. Yang, M. Chen, W. Saad, W. Xu, and M. Shikh-Bahaei, "Sum-rate maximization of uplink rate splitting multiple access (RSMA) communication," *IEEE Transactions on Mobile Computing*, vol. 21, no. 7, pp. 2596–2609, 2022.

[20] M. Diamanti, C. Pelekis, E. E. Tsiropoulou, and S. Papavassiliou, "Delay minimization for rate-splitting multiple access-based multi-server MEC offloading," *IEEE/ACM Transactions on Networking*, pp. 1–13, 2023.

[21] T. P. Truong, N.-N. Dao, and S. Cho, "HAMEC-RSMA: Enhanced aerial computing systems with rate splitting multiple access," *IEEE Access*, vol. 10, pp. 52 398–52 409, 2022.

[22] P. Chen, H. Liu, Y. Ye, L. Yang, K. J. Kim, and T. A. Tsiftsis, "Rate-splitting multiple access aided mobile edge computing with randomly deployed users," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 5, pp. 1549–1565, 2023.

[23] L. N. Ribeiro, S. Schwarz, M. Rupp, and A. L. F. de Almeida, "Energy efficiency of mmWave massive MIMO precoding with low-resolution DACs," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 2, pp. 298–312, 2018.

[24] J. Zhang, L. Dai, Z. He, B. Ai, and O. A. Dobre, "Mixed-ADC/DAC multipair massive MIMO relaying systems: Performance analysis and power optimization," *IEEE Transactions on Communications*, vol. 67, no. 1, pp. 140–153, 2019.

[25] J. Dai, Y. Wang, C. Pan, K. Zhi, H. Ren, and K. Wang, "Reconfigurable intelligent surface aided massive MIMO systems with low-resolution DACs," *IEEE Communications Letters*, vol. 25, no. 9, pp. 3124–3128, 2021.

[26] R. Wang, H. Ren, C. Pan, J. Fang, M. Dong, and O. A. Dobre, "Channel estimation for RIS-aided mmWave massive MIMO system using few-bit ADCs," *IEEE Communications Letters*, vol. 27, no. 3, pp. 961–965, 2023.

[27] X. Li, M. Zhang, H. Chen, C. Han, L. Li, D.-T. Do, S. Mumtaz, and A. Nallanathan, "UAV-enabled multi-pair massive MIMO-NOMA relay systems with low-resolution ADCs/DACs," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 2, pp. 2171–2186, 2024.

[28] S. Kim, J. Choi, and J. Park, "Downlink NOMA for short-packet internet of things communications with low-resolution ADCs," *IEEE Internet of Things Journal*, vol. 10, no. 7, pp. 6126–6139, 2023.

[29] S. Park, J. Choi, J. Park, W. Shin, and B. Clerckx, "Rate-splitting multiple access for quantized multiuser MIMO communications," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2023.

[30] M. Katwe, K. Singh, B. Clerckx, and C.-P. Li, "Rate splitting multiple access for sum-rate maximization in IRS aided uplink communications," *IEEE Transactions on Wireless Communications*, vol. 22, no. 4, pp. 2246–2261, 2023.

[31] R. Zhang, K. Xiong, Y. Lu, P. Fan, D. W. K. Ng, and K. B. Letaief, "Energy efficiency maximization in RIS-assisted SWIPT networks with RSMA: A PPO-based approach," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 5, pp. 1413–1430, 2023.

[32] H. Niu, Z. Lin, K. An, J. Wang, G. Zheng, N. Al-Dhahir, and K.-K. Wong, "Active RIS assisted rate-splitting multiple access network: Spectral and energy efficiency tradeoff," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 5, pp. 1452–1467, 2023.

[33] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 4157–4170, 2019.

[34] J. Zhao, L. Yu, K. Cai, Y. Zhu, and Z. Han, "RIS-aided ground-aerial NOMA communications: A distributionally robust DRL approach," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 4, pp. 1287–1301, 2022.

[35] H. Yu, H. D. Tuan, E. Dutkiewicz, H. V. Poor, and L. Hanzo, "RIS-aided zero-forcing and regularized zero-forcing beamforming in integrated information and energy delivery," *IEEE Transactions on Wireless Communications*, vol. 21, no. 7, pp. 5500–5513, 2022.

[36] Y. Guo, P. Sun, Z. Yuan, C. Huang, Q. Guo, Z. Wang, and C. Yuen, "Efficient channel estimation for RIS-aided MIMO communications with unitary approximate message passing," *IEEE Transactions on Wireless Communications*, vol. 22, no. 2, pp. 1403–1416, 2023.

[37] A. K. Fletcher, S. Rangan, V. K. Goyal, and K. Ramchandran, "Robust predictive quantization: Analysis and design via convex optimization," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 618–632, 2007.

[38] J. Choi, Y. Cho, and B. L. Evans, "Quantized massive MIMO systems with multicell coordinated beamforming and power control," *IEEE Transactions on Communications*, vol. 69, no. 2, pp. 946–961, 2021.

[39] G. Chen, Q. Wu, C. He, W. Chen, J. Tang, and S. Jin, "Active IRS aided multiple access for energy-constrained IoT systems," *IEEE Transactions on Wireless Communications*, vol. 22, no. 3, pp. 1677–1694, 2023.

[40] S. Cui, A. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2349–2360, 2005.

[41] R. Long, Y.-C. Liang, Y. Pei, and E. G. Larsson, "Active reconfigurable intelligent surface-aided wireless communications," *IEEE Transactions on Wireless Communications*, vol. 20, no. 8, pp. 4962–4975, 2021.

[42] S. Liu, Y. Yu, X. Lian, Y. Feng, C. She, P. L. Yeoh, L. Guo, B. Vucetic, and Y. Li, "Dependent task scheduling and offloading for minimizing deadline violation ratio in mobile edge computing networks," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 2, pp. 538–554, 2023.

[43] Z. Gu, C. She, W. Hardjawana, S. Lumb, D. McKechnie, T. Essery, and B. Vucetic, "Knowledge-assisted deep reinforcement learning in 5g scheduler design: From theoretical framework to implementation," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 7, pp. 2014–2028, 2021.

[44] J. Xu, B. Ai, and T. Q. S. Quek, "Toward interference suppression: RIS-aided high-speed railway networks via deep reinforcement learning," *IEEE Transactions on Wireless Communications*, vol. 22, no. 6, pp. 4188–4201, 2023.

[45] H. Wang, H. Zhang, X. Liu, K. Long, and A. Nallanathan, "Joint UAV placement optimization, resource allocation, and computation offloading for THz band: A DRL approach," *IEEE Transactions on Wireless Communications*, vol. 22, no. 7, pp. 4890–4900, 2023.

[46] K. K. Nguyen, A. Masaracchia, V. Sharma, H. V. Poor, and T. Q. Duong, "RIS-assisted UAV communications for IoT with wireless power transfer using deep reinforcement learning," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 5, pp. 1086–1096, 2022.

[47] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," 2019.

[48] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski

et al., "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.
[49] G. E. Uhlenbeck and L. S. Ornstein, "On the theory of the brownian motion," *Physical review*, vol. 36, no. 5, p. 823, 1930.
[50] V.-D. Nguyen, H. V. Nguyen, O. A. Dobre, and O.-S. Shin, "A new design paradigm for secure full-duplex multiuser systems," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 7, pp. 1480–1498, 2018.
[51] V. Kumar, R. Zhang, M. D. Renzo, and L.-N. Tran, "A novel SCA-based method for beamforming optimization in IRS/RIS-assisted MU-MISO downlink," *IEEE Wireless Communications Letters*, vol. 12, no. 2, pp. 297–301, 2023.
[52] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," https://cvxr.com/cvx, Mar. 2014.
[53] ——, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds. Springer-Verlag Limited, 2008, pp. 95–110, http://stanford.edu/∼boyd/graph_dcp.html.
[54] T. P. Truong, V. D. Tuong, N.-N. Dao, and S. Cho, "FlyReflect: Joint flying IRS trajectory and phase shift design using deep reinforcement learning," *IEEE Internet of Things Journal*, vol. 10, no. 5, pp. 4605–4620, 2023.
[55] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
[56] D. Gunasinghe and G. A. A. Baduge, "Achievable rate analysis for multi-cell RIS-aided massive MIMO with statistical CSI-based optimizations," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2023.
[57] Z. Wei, Y. Cai, Z. Sun, D. W. K. Ng, J. Yuan, M. Zhou, and L. Sun, "Sum-rate maximization for IRS-assisted UAV OFDMA communication systems," *IEEE Transactions on Wireless Communications*, vol. 20, no. 4, pp. 2530–2550, 2021.
[58] X. Ma, Z. Chen, W. Chen, Z. Li, Y. Chi, C. Han, and S. Li, "Joint channel estimation and data rate maximization for intelligent reflecting surface assisted terahertz MIMO communication systems," *IEEE Access*, vol. 8, pp. 99 565–99 581, 2020.