Energy-Efficient Communication Networks via Multiple Aerial Reconfigurable Intelligent Surfaces: DRL and Optimization Approach

Pyae Sone Aung, Yu Min Park, Yan Kyaw Tun, Zhu Han, Fellow, IEEE, and Choong Seon Hong, Senior Member, IEEE

Abstract—In the realm of wireless communications in 5G, 6G and beyond, deploying unmanned aerial vehicle (UAV) has been an innovative approach to extend the coverage area due to its easy deployment. Moreover, reconfigurable intelligent surface (RIS) has also emerged as a new paradigm with the goals of enhancing the average sum-rate as well as energy efficiency. By combining these attractive features, an energy-efficient RIS-mounted multiple UAVs (aerial RISs: ARISs) assisted downlink communication system is studied. The novelty of this research lies in the joint optimization problem of deployment of ARIS, ARIS reflecting elements on/off states, phase shift, and power control of the multiple ARIS-assisted communication system is formulated. The problem is challenging to solve since it is mixed-integer, non-convex, and NP-hard. To overcome this, it is decomposed into three sub-problems. Afterward, successive convex approximation (SCA), actor-critic proximal policy optimization (AC-PPO), and whale optimization algorithm (WOA) are employed to solve these sub-problems alternatingly. Finally, extensive simulation results have been generated to illustrate the efficacy of our proposed algorithms.

Index Terms—Aerial reconfigurable intelligent surface (ARIS), RIS, unmanned aerial vehicle (UAV), deployment, reflecting elements on/off states, phase shift, transmit power optimization, successive convex optimization (SCA), actor-critic proximal policy optimization (AC-PPO), whale optimization algorithm (WOA).

I. INTRODUCTION

A. Background and Motivations

As claimed by Cisco Networking Index (CNI), the number of Internet users reached 3.9 billion in 2018 and is anticipated to surpass 5.3 billion by 2023 [1]. The rapid growth of multimedia devices such as the Internet of Things (IoT), video streaming, online gaming, Virtual Reality (VR), and Augmented Reality (AR) applications, thrives immense challenges for current communication architecture and motivates to discover new ways to enhance spectral efficiency in both academic and industrial fields. Numerous ingenious wireless technologies have been developed in the last several years, which include deploying unmanned aerial vehicles (UAVs) and Reconfigurable Intelligent Surfaces (RIS) elements.

Recently, UAVs have achieved a great deal of interest in deploying as a communication and computing platform due to their high mobility and ease of deployment. The emplacement of UAVs can not only save the cost of mobile infrastructure which demands a large budget but also save time for quick on-demand deployment to provide services in rural regions or disaster areas or temporary events such as concerts, stadiums where the infrastructure is difficult to come across. In some scenarios [2], [3], UAVs are implemented with a multi-access edge computing (MEC) system to deliver the computing resources near to the user equipment (UE) which saves a considerably large amount of time for uploading, computing and downloading tasks.

The newly recent technology called RIS, which is incited from the recent development of meta-surfaces, benefits the wireless communications in extending the coverage range and improving the signal quality at the receiver [4]. RIS is a man-made meta-surface implemented with low-cost passive elements that can be programmed by integrated electronic circuits to alter the incoming electromagnetic field in a desirable way [5]. Unlike the traditional collaborative communications such as decode-and-forward (DF) and amplify-and-forward (AF), RIS does not require an additional power amplifier hence, is more environmentally friendly and energy-efficient [6]. Taking into account of its cost effectiveness and energy efficiency, RIS technology has acquired vast attention in 5G, 6G, and beyond communications. Furthermore, since RIS structures consist of relatively small hardware components, they can be easily integrated into several...
communication environments, such as along the surfaces of the building [7].

The simultaneous deployment of RIS mounted on UAV, also known as aerial RIS (ARIS), offers several benefits. One of the significant advantages of ARIS is enhanced coverage and capacity. A UAV can quickly move to different locations and altitudes, while RIS can reflect the signals and manipulate the signal propagation path to provide enhanced coverage and signal quality [8]. The deployment of ARIS can offer mobility and flexibility, since it can be deployed in areas that are difficult to reach with traditional wireless communication systems [9]. Furthermore, the use of ARIS can improve the energy efficiency of wireless communication systems. The RISs can reflect the incident signals with phase shifts that constructively add up at the receiver, thereby enhancing the signal-to-noise ratio (SNR) and reducing the required transmission power levels. By optimizing the deployment and phase shifts of the RIS elements, the signal quality can be improved while minimizing the power consumption [10]. In addition, ARIS can enhance the security of wireless networks by reflecting the signals away from unauthorized receivers and directing them to specific receivers. This can offer a more secure and private wireless communication system [11].

B. Challenges and Research Contributions

When UAVs are considered as communication and computing platforms, there exist several challenges in UAVs’ energy consumption as they are energy-constrained devices. Additionally, even though RIS can enhance the spectral efficiency, setting up RIS structures to achieve Line-of-sight (LoS) links between UE and RIS is still a quite challenging issue. Taking the advantages of RIS in enhancing spectral efficiency without the requirement of any external power sources with the aid of UAVs to obtain LoS links between UE and RIS, we propose an energy-efficient multiple ARISs-assisted system to extend the downlink communication links from the ground base station (BS) to the UEs. In our system, we assume that there is no dominant LoS links between the BS and UEs due to obstacles. Several studies have examined merely improving the downlink data rate while ignoring energy efficiency [12], [13]. The main idea to enhance our energy-efficient multiple ARISs-assisted networks is that while the power of a single RIS on a UAV typically consumes less than the propulsion system, the energy consumption increases for several ARISs. By maximizing energy efficiency while taking into account the deployment, on/off states of reflecting elements, phase shifts, and power control for the multiple ARISs, we can conserve the significant energy of the whole system. The contributions of this paper can be organized as follows:

- Firstly, we propose the downlink communication system between the BS and UEs, which is assisted by the multiple ARISs to enhance the spectral efficiency for all UEs since the dominant LoS link between BS and UEs are blocked by the obstacles. We assume the BS and ARISs are deployed by the same service operator, and thus the BS is responsible for ARIS deployment and controlling the on/off states and the phase shifts for the ARIS reflecting elements.
- Secondly, we formulate the problem to maximize the energy efficiency of the proposed system by jointly optimizing the ARISs deployment, ARIS reflecting elements on/off states, phase shift, and power control. We show that the formulated problem is a mixed integer non-linear programming (MINLP) problem, and it is challenging to solve in polynomial time.
- To address this challenge, we decompose our formulated problem into three sub-problems: 1) ARISs deployment problem, 2) joint ARIS reflecting elements on/off states and phase shift problem, and 3) power control problem. Then, successive convex approximation (SCA), actor-critic proximal policy optimization (AC-PPO), and whale optimization algorithm (WOA) are proposed to solve these sub-problems alternatingly.
- Finally, a comprehensive numerical analysis is integrated to validate the efficacy of the overall performance of our proposed algorithms with several benchmark schemes, such as single-ARIS, ARIS with fixed phase shifts (ARIS-NPS), and UAV as relay (UAV-relay) scenarios. We achieve an improvement in average sum-rate by 18% and 51% compared to the single-ARIS and the ARIS-NPS scenarios, and 30% and 68% increase in energy efficiency compared to the single-ARIS and the UAV-relay scenarios, respectively. Moreover, our proposed multiple ARISs-assisted system achieves 33% increase in average sum-rate compared to the multiple RISs-assisted system.

The rest of the article is categorized as follows: we present the related works in Section II. Next, we present our system model and problem formulation in Section III. Afterwards, the solution approach is proposed in Section IV, and performance evaluation is performed in Section V. Finally, Section VI concludes our article.

II. RELATED WORKS

A. UAV-Assisted Wireless Networks

An overview on the literature related to UAV-assisted wireless networks is discussed in this section [14], [15], [16], [17], [18], [19], [20]. The major strength of UAV in enhancing coverage area, energy efficiency, and cost efficiency has received significant attention in recent years [14]. In [15], the authors studied to maximize the uplink communication where UAVs served as relays. In [16], the authors studied a single UAV-assisted device-to-device (D2D) communications and analysed how the UAV’s altitude can impact the rate performance and coverage area on the D2D users’ density. The authors in [17] derived the channel model for the LoS air-to-ground UAV communications. There exist several works that studied the UAV deployment [18], [19], [20], [21]. For instance, the authors in [18] studied the UAVs deployment for UAV-to-ground communication in arbitrary spatial distribution for network planning to provide wireless services to the ground users and the authors in [19] studied the incorporation between UAVs in a 3D cellular network. The work in [20] studied the adaptive UAV deployment for dynamic users. The authors in [21] studied DRL-based dynamic UAV control instead of static UAV deployment. In all of the above works, UAV is
considered either as aerial BSs or MEC devices or relays, which results in higher energy consumption.

B. RIS-Assisted Wireless Networks

An overview on the literature related to RIS-assisted wireless networks are discussed in this section [5], [22], [23], [24], [25], [26], [27], [28], [29]. In [5], the authors considered developing an energy-efficient architecture for the RIS structures in accordance with power allocation and phase shifting values of RIS elements while guaranteeing the individual data rate budget for each user. In [22], the authors proposed the energy-harvesting RIS elements implemented on the facades of the buildings in order to maximize the spectral efficiency while enabling the transmit power control and RIS configuration under the indeterminate wireless channel condition. The authors in [23] aimed to distinguish the principal relationship between the total sum-rate of multiple users and the required number of RIS reflecting elements in wireless communications. They observed that the capacity of the system could no longer efficiently rise as the number of RIS elements reached the upper bound limit. They also investigated how the number of phase shifts can effect the performance on the achievable data rate. The authors in [24] investigated the practical case study between phase shift and finite-sized RIS to maximize the downlink multi-user system. In [25], the authors studied about RIS elements to eliminate interference between multiple D2D uplink communication networks. There have also been several studies on RIS-assisted in the vehicular networks. In [26], the authors investigated the secrecy outage probability of vehicular-to-vehicular (V2V) and vehicular-to-infrastructure (V2I) communications. The authors in [27] aimed to maximize the data rate of each vehicle where the communication links from the road site unit (RSU) is extended by the RIS technology with discrete phase shift. The authors in [28] studied deep reinforcement learning (DRL)-based RIS-assisted multi-user downlink multiple input single output (MISO) system. The work in [29] considered improving the secrecy rate of users in RIS-assisted system by constructing a DRL-based QoS-aware reward function. All of the aforementioned works only considered the RIS-assisted networks, where RIS elements are either implemented on the ground level or facades of the building, which is still challenging to achieve the dominant LoS communication links between the BS-RIS-users.

C. UAV-RIS-Assisted Wireless Networks

The idea of RIS installed on UAV has received attention due to the reception of strong LoS communication in recent years. An overview on the literature related to UAV-RIS-assisted wireless networks is discussed in this section [9], [10], [30], [31], [32], [33], [34], [35], [36], [37], [38]. The authors in [30] examined the adaptive RIS-assisted aerial-terrestrial downlink communication system between UAVs and multi-users with respect to RIS elements allocation and reflecting coefficients. In [31], the authors looked into UAV-user communication with RIS assistance in order to maximize the worst-case secrecy rate by taking into account of the transmitter’s power allocation, RIS’s beamforming and UAV’s trajectory. The authors in [32] proposed the RIS-assisted UAV communications to maximize the received signal power at the ground user by considering the passive and active beamforming and UAV’s trajectory. Furthermore, in [33], the authors minimized the energy consumption for both orthogonal multiple access (OMA) and non-orthogonal multiple access (NOMA) cases by jointly considering the trajectory for the UAV and passive beamforming of the RIS elements. There also exist several works on ARIS-assisted system [9], [34]. In [35], the authors considered ARIS-assisted system to satisfy the constraints of ultra-reliable low latency communication (URLLC). The authors in [36] studied the several UAVs-RISs-assisted total transmit power minimization for the heterogeneous networks. However, they overlooked the energy efficiency of the system. By jointly controlling the BS selection and beamforming, the authors in [37] considered maximizing energy efficiency for a single ARIS-assisted downlink communication for a single user for disaster relief. The authors in [10] maximized the energy efficiency of the covert communication system with the aid of a single ARIS. In addition, in [38], the authors considered implementing an ARIS as a passive relay node in intelligent IoT networks to assist IoT devices to transfer the sampled data to the BS. All of the aforementioned works only evaluated a single ARIS which prevents them from taking full advantage of beamforming gain. To tackle this issue, there also exist several studies considering multiple aerial RISs-assisted communication frameworks [8], [12], [13], [36]. The authors in [8] studied the network coverage extension for massive MIMO communication networks with the assistant of multiple ARISs. Likewise, the authors in [12] and [13] examined and analysed to improve the downlink communication of multi UAV-RISs and swarm-enabled ARISs assisted wireless networks with joint beamforming design, phase-shift and placement of ARISs. However, the abovementioned works did not account for the on/off states of reflecting elements which could preserve more power for energy-constrained ARISs. In this article, we propose the multiple ARISs in order to maximize the average energy efficiency for the downlink communication between the BS and the UEs by jointly optimizing the ARISs deployment, ARIS reflecting elements on/off state, phase shift, and power control, which significantly impacts the performance of the system.

III. SYSTEM MODEL

Our system model includes a BS $B$ with multiple antennas, a set $\mathcal{N}$ of $N$ ARISs in which each RIS is implemented on each UAV, and each ARIS, $n \in \mathcal{N}$, contains an array of $I_n = \{I_{n_1}, I_{n_2}, \ldots, I_{n_l}\}$ reflecting elements and a set $\mathcal{K}$ of $K$ UEs with a single antenna as shown in Fig. 1. To reduce air-interface latency, we assume that the ARISs are operated by the same service provider at the BS, and implemented with low-latency wireless communication protocol such as IEEE 802.11ax [39], IEEE 802.11be [40]. The coordinates of the BS is denoted by $q_B = (x_B, y_B, z_B)$, where $z_B$ is the height of the BS. Similarly, the positions of each UE $k$ and each ARIS $n$ can be represented as $q_k = (x_k, y_k, 0)$ and $q_n = (x_n, y_n, z_n)$ respectively, and $z_n$ is the height where the RIS-mounted UAV is hovering. The
time horizon of the system can be divided into a discrete set of $T = \{1, 2, \ldots, t, \ldots, T\}$.

Since ARISs have limited energy, apart from hovering, the ARIS reflecting elements need to be turned off when there is no active users in order to reserve excessive energy. In our work, we assume that the reflecting elements of ARIS are active, and we acknowledge that dynamically controlling the activation of reflecting elements based on the signal strength between the BS, ARISs, and UEs can lead to significant energy savings. Authors in [41] and [30] proved that turning off the whole RIS or some surface area of RIS can preserve energy. The on/off states of reflecting element $i$ in each ARIS $n$ are controlled by the decision variable $\delta_{in}$ as follows:

$$\delta_{in}[t] = \begin{cases} 1, & \text{if reflecting element } i \text{ of ARIS } n \text{ is switched on at time } t, \\ 0, & \text{otherwise}. \end{cases}$$ (1)

Next, we define $\Delta_n[t] = \text{diag}\{\delta_{in}[t]\} \in \mathbb{R}^{I_n \times I_n}$ as the diagonal on/off states matrix for all reflecting elements $I_n$ for each ARIS $n$ to decide whether to turn on or off, where $\delta_{in}[t] = [\delta_{i_1n}[t], \delta_{i_2n}[t], \ldots, \delta_{i_{I_n}n}[t]]$.

A. Communication Model

We assume that perfect channel state information (CSI) is known at the BS, and CSI acquisition stay out of the scope of this article, but can be found in [23], [42]. We adopt both direct and indirect communication links between the BS and the UEs. For the direct link, we assume there is no dominant propagation along the LoS signal between the BS and UEs. Therefore, we adopt the Rayleigh fading model and the channel gain for the BS-UE link at time $t$ can be obtained as follows:

$$h_{B,n}[t] = \sqrt{\kappa d_{B,n}^{-\alpha}} \|q_{B,t} - q_{k,t}\|,$$ (2)

where $\kappa$ is the channel gain at the reference distance 1 m, $\alpha \geq 2$ is the path loss exponent, $|d_{B,k}[t]| = |q_{B,t} - q_{k,t}|$ is the Euclidean distance between the BS and UE $k$ at time $t$, and $\hat{h}$ is the complex Gaussian random scattering component with zero mean and unit variance.

For the indirect communication, there exist two links: BS-ARIS link and ARIS-UE link, respectively. For the BS-ARIS link, we consider there is only LoS signal between the BS and ARIS, and thus the channel fading here is assumed to experience the Rician channel fading with only LoS components. Therefore, the channel gain between the BS and ARIS $n$ at time $t$ can be defined as:

$$H_{B,n}[t] = \sqrt{\kappa d_{B,n}^{-\alpha}} \|q_{B,t} - q_{k,t}\| \frac{R}{1 + R} H_{\text{LoS},B,n}[t],$$ (3)

where $\hat{R}$ is the Rician factor, and $|d_{B,n}[t]| = |q_{n,t} - q_{k,t}|$ is the distance between the BS to ARIS $n$ at time $t$. $H_{\text{LoS},B,n}[t]$ is the deterministic LoS component between the BS and ARIS $n$ in correspondence with the azimuth angle-of-arrival (AoA) of the link at time $t$ [30]. For the ARIS-UE link, there are both LoS and non-line-of-sight (NLoS) propagation between ARISs and UEs. Consequently, the Rician fading model is adopted and the channel gain for the ARIS-UE link at time $t$ can be obtained as follows:

$$h_{n,k}[t] = \sqrt{\kappa d_{n,k}^{-\alpha}} \|q_{n,t} - q_{k,t}\| \frac{R}{1 + R} h_{\text{LoS},n,k}[t] + \frac{1}{1 + R} h_{\text{NLOS},n,k},$$ (4)

where $|d_{n,k}[t]| = |q_{n,t} - q_{k,t}|$ is the distance between ARIS $n$ and UE $k$ at time $t$. $h_{\text{LoS},n,k}[t]$ is the deterministic LoS component between ARIS $n$ and UE $k$ corresponding with the azimuth angle-of-departure (AoD) of the link, and $h_{\text{NLOS},n,k}$ is the non-LoS component, which follows the identically and independently distributed circularly-symmetric complex Gaussian distribution.

Furthermore, at time $t$, the incident signals are reflected by each reflecting element $i$ of ARIS $n$ from the feasible range of phase shift values specified by

$$\theta_{in}[t] = e^{(2\pi b + \phi)},$$ (5)

where $\phi$ is the phase shift index, and $b$ is the phase shift resolution in bits [43]. Therefore, a vector of $\Theta_{in}[t] = [\theta_{i_1n}[t], \theta_{i_2n}[t], \ldots, \theta_{i_{I_n}n}[t]]$ represents the phase shift values of ARIS $n$. Following that, the reflection coefficient matrix can be denoted by

$$\Theta_{n}[t] = \text{diag}(\beta_{1n} e^{j\theta_{1n}[t]}, \beta_{2n} e^{j\theta_{2n}[t]}, \ldots, \beta_{I_n} e^{j\theta_{I_n}n[t]}),$$ (6)

where $\beta_{in} \in [0, 1]$ denotes the amplitude reflection coefficient of the $i$-th reflecting element of the $n$-th ARIS, and $j$ is the imaginary unit of a complex number. Therefore, the received signal at UE $k$ can be achieved as follows:

$$y_{k}[t] = (h_{B,k}[t] + \sum_{n=1}^{N} h_{n,k}[t] \Delta_{n}[t] \Theta_{n}[t] H_{B,n}[t]) x + \omega_{k},$$ (7)

where $x = \sum_{k=1}^{K} g_{k[t]} s_{k}$ is the transmitted signal from the BS with beamforming vector $g_{k[t]}$ at time $t$, and the unit-power complex based information symbol $s_{k}$ for UE $k$, while $\omega_{k} \sim CN(0, \sigma^2)$ denotes the additive white Gaussian noise (AWGN) at UE $k$. Based on (7), the signal-to-interference-plus-noise ratio...
can be obtained as
\begin{equation}
\gamma_k[t] = \frac{1}{2} \left[ \frac{1}{\alpha} \sum_{n=1}^{N} h_{n,k}[t] \Delta_n[t] \Theta_n[t] H_{B,n}[t] g_k[t] \right]^2 + \sigma^2.
\end{equation}

Afterwards, based on (8), the achievable data rate of UE \( k \) can be formulated as follows:
\begin{equation}
r_k[t] = W \log_2 (1 + \gamma_k[t]),
\end{equation}
where \( W \) is the transmission bandwidth available for each UE. Therefore, the sum-rate of all users can be described as follows:
\begin{equation}
R[t] = \sum_{k=1}^{K} r_k[t].
\end{equation}

B. Power Consumption Model

In our scenario, we need to consider the propulsion’s power consumption to support the movement of ARIS if necessary. Since we assume the rotary-wing UAVs with constant speed \( V \), the propulsion power consumption for each UAV \( n \) can be obtained as follows [44]:
\begin{align}
P_{UAV}^n & = P_0 \left( 1 + \frac{3 V^2}{U_{tip}^2} \right) + P_i \left( \sqrt{1 + \frac{V^4}{4 v_0^4}} - \frac{V^2}{2 v_0^2} \right)^{1/2} \\
& + \frac{1}{2} d_0 \phi \Delta \eta V^3,
\end{align}
where \( P_0 = \frac{\psi \varphi \Delta \eta \rho V^3}{\phi} \) is the power required to rotate the rotor blades with \( \psi, \varphi, \Delta, \eta, v_0, \) and \( \phi \) representing the coefficient of the profile drag, density of the air, rotor solidity, disc area of the rotor, blade angular velocity, and radius of the rotor, respectively. Next, \( P_i = (1 + \iota \frac{D(t)}{\sqrt{\phi/n^3}}) \) is the power required to endure the induced drag generated by the lift, with \( \iota \) and \( \phi \) denoting the incremental correction factor, and weight of the aircraft, respectively. Moreover, \( U_{tip} \) is the tip speed of the rotor blade, \( v_0 \) is the mean rotor induced velocity in hover, and \( d_0 \) is fuselage drag ratio, respectively.\(^1\)

Furthermore, in this work, the BS controls the phase shifts of the ARIS reflecting elements. Hence, the total power of the considered multiple ARISs-assisted downlink system includes [41]:
1) transmit power of the BS, 2) circuit power of each UE \( k \), 3) circuit power of BS 4) circuit power consumption of ARIS and 5) power consumption of the rotary-wing UAV, and is defined as
\begin{equation}
P[t] = \sum_{k=1}^{K} (\zeta g_k[t] H g_k[t] + P_{AR}^n) + P_B \\
+ \sum_{n=1}^{N} (\Delta_n[t] P_{n,AR} + P_{UAV}^n),
\end{equation}
where \( \zeta = 1/\mu \) with \( \mu \) being the transmit power amplifier efficiency, \( P_{AR}^n \) is the circuit power of each user \( k \), \( P_B \) is the circuit power of BS, and \( P_{n,AR} \) is the circuit power consumption of ARIS \( n \). The transmit signal power of the BS has the constraint as follows:
\begin{equation}
tr(g[t] H g[t]) \leq P_{max}, \forall t \in T,
\end{equation}
where \( tr(S) \) means the trace of square matrix \( S, g = [g_1; \ldots; g_K] \) and \( P_{max} \) is the maximum transmission power available at the BS.

C. Problem Formulation

The main objective of this work is to maximize energy efficiency of the system, i.e., to maximize the average sum-rate \( R[t] \) for the UEs under the constraint of the power consumption \( P[t] \) of both ARISs and the BS. To accomplish this, we need to jointly optimize the deployment of ARIS, ARIS reflecting element on/off states, phase shift, and power control of the BS. Prior to problem formulation, we define the required constraints as follows:

Each UE \( k \) is necessary to fulfill the demand for the specified data rate at time \( t \), which is defined as:
\begin{equation}
r_k[t] \geq r_k^{\min}[t], \forall k \in K, \forall n \in N', \forall i \in I_n, \forall t \in T,
\end{equation}
where \( r_k^{\min}[t] \) is the minimum data rate requirement of UE \( k \) at time \( t \). The accessible phase shift value of \( i \)-th reflecting element \( n \)-th ARIS at time \( t \) should be between 0 to 2\( \pi \) as follows:
\begin{equation}
0 \leq \theta_{in}[t] < 2\pi, \forall n \in N', \forall i \in I_n, \forall t \in T.
\end{equation}
A safe distance between two adjacent ARISs is necessary to ensure collision avoidance and interference caused by other ARIS [45], [46]. We denote \( d_{min} \) as the threshold distance between two adjacent ARISs at time \( t \), and can be defined as follows:
\begin{equation}
||q_i[t] - q_j[t]|| \geq d_{min}, \forall i \in N', \forall j \in N', i \neq j, \forall t \in T.
\end{equation}
Furthermore, each reflecting element \( i \) of ARIS \( n \) can only be either turned on or off at one time slot and is given as follows:
\begin{equation}
\delta_{in}[t] \in \{0, 1\}, \forall n \in N', \forall i \in I_n, \forall t \in T.
\end{equation}

Given the above mentioned network characteristics, our optimization problem \( E \) can be mathematically formulated as follows:
\begin{equation}
P : \quad \max_{q, \Delta, \Theta, g} E(q, \Delta, \Theta, g)
\end{equation}
s.t.
\begin{align}
r_k[t] & \geq r_k^{\min}[t], \forall k \in K, \forall n \in N', \forall i \in I_n, \forall t \in T \\
0 & \leq \theta_{in}[t] < 2\pi, \forall n \in N', \forall i \in I_n, \forall t \in T \\
||q_i[t] - q_j[t]|| & \geq d_{min}, \forall i \in N', \forall j \in N', i \neq j, \forall t \in T \\
tr(g[t] H g[t]) & \leq P_{max}, \forall t \in T \\
\delta_{in}[t] & \in \{0, 1\}, \forall n \in N', \forall i \in I_n, \forall t \in T.
\end{align}

\(^1\)For the sake of simplicity, we disregard the aircraft’s weight shift when fuel or batteries are depleted over time.
Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
Initially, we define the first-order Taylor series as a linear approximation of the function. To do so, we utilize the first-order Taylor expansion to find the linear approximation of the function. To do so, we can transform the first-order Taylor expansion to replace the \( \hat{l}(\bar{r}) \), we can construct its lower bound as follows:

\[
\hat{E}(\bar{r}, \bar{r}') = \hat{h}(\bar{r}) - \hat{\bar{I}}((\bar{r}, \bar{r}')), 
\]

where

\[
\hat{l}((\bar{r}, \bar{r}')) = \frac{\partial \hat{\bar{I}}((\bar{r}, \bar{r}'))}{\partial \bar{r}'},
\]

and

\[
\hat{l}(\bar{r}) = \frac{\partial \hat{l}(\bar{r})}{\partial \bar{r}} = \frac{1}{\ln 2} \sum_{k=1}^{K} \left( h_{B,k}[t] + \sum_{n=1}^{N} \kappa g_n[k][\bar{r}][t] + \sigma^2 \right).
\]

Both the functions \( \hat{h}(\bar{r}) \) and \( \hat{l}(\bar{r}) \) are convex. However, the difference between them is neither convex nor concave, as represented in (27). Then, we find the feasible solution \( \bar{r}' \) to problem (23) by computing the concave lower bound, i.e., the surrogate function of the non-concave objective, specified in (27). By implementing

\[
\hat{E}(\bar{r}) = \sum_{k=1}^{K} W \log_2 \left( 1 + \frac{\hat{h}_{B,k}[t] + \sum_{n=1}^{N} \kappa g_n[k][\bar{r}][t]}{\sum_{l=1, l \neq k} \hat{h}_{B,l}[t] + \sum_{n=1}^{N} \kappa g_n[l][\bar{r}][t] + \sigma^2} \right).
\]
Afterwards, we have
\[ \hat{t}(\hat{p}) \leq \hat{t}(p') + \nabla \hat{t}(p')(\hat{p} - p'). \] (34)

Therefore, we can derive the observations by examining (27), (30), and (34) as follows:
\[ \hat{E}(\hat{p}) = \hat{h}(\hat{p}) - \hat{t}(\hat{p}) \]
\[ \geq \hat{h}(\hat{p}) - \{ \hat{t}(p') + \nabla \hat{t}(p')(\hat{p} - p') \} \]
\[ = \hat{E}(\hat{p}, p'), \] (35)

where (35) denotes that the surrogate function provides the lower bound of the original function. As a result, at point \( \hat{p}' \), i.e., \( \hat{E}(\hat{p}, \hat{p}') |_{\hat{p}=\hat{p}'} = \hat{E}(\hat{p}') \), the two functions are tangent to each other. Thereby, our objective function of sub-problem (23) has the lower bound function as obtained in (35).

Consequently, we replace our objective function in sub-problem (23) which is non-convex, by its surrogates as presented in (30). Furthermore, we take the first-order Taylor expansions of \( (a[t])^{-\frac{1}{\alpha}} \), \( (b[t])^{-\frac{1}{\beta}} \), and \( H_{ab}[t]H'[t]H_{ab}[t] \) at the given feasible points \( a_0 = \{ a_0[i] \}_{i=1}^T \), \( b_0 = \{ b_0[i] \}_{i=1}^T \), and \( H_{0ab} = \{ H_{0ab}[t] \}_{i=1}^T \) are expressed as follows:
\[ (a[t])^{-\frac{1}{\alpha}} \geq (a_0[t])^{-\frac{1}{\alpha}} - \frac{4}{\alpha} (a_0[t])^{-\frac{1}{\alpha} - 1} a[t] \] (36)
\[ (b[t])^{-\frac{1}{\beta}} \geq (b_0[t])^{-\frac{1}{\beta}} - \frac{4}{\beta} (b_0[t])^{-\frac{1}{\beta} - 1} b[t] \] (37)
\[ H_{0ab}[t]H'[t]H_{0ab}[t] \geq -H_{0ab}[t]H'[t]H_{0ab}[t] \]
\[ + 2R \left[ H_{0ab}[t]H'[t]H_{0ab}[t] \right]. \] (38)

By combining (25) and (36), (26) and (37), we get
\[ x_B^2 + x_B[t]^2 + y_B^2 + y_B[t]^2 - 2x_Bx_B[t] - 2y_By_B[t] \]
\[ + (z_B - z_B[t])^2 - \left( 1 + \frac{4}{\alpha} \right) (a_0[t])^{-\frac{1}{\alpha}} - \frac{4}{\alpha} (a_0[t])^{-\frac{1}{\alpha} - 1} a[t] \leq 0, \] (39)
\[ x_n[t]^2 + x_n[t]^2 + y_n[t]^2 + y_n[t]^2 - 2x_n[x_n[t]x_n[t] - 2y_n[y_n[t]] \]
\[ + (z_n - z_n[t])^2 - \left( 1 + \frac{4}{\alpha} \right) (b_0[t])^{-\frac{1}{\beta}} - \frac{4}{\beta} (b_0[t])^{-\frac{1}{\beta} - 1} b[t] \leq 0. \] (40)

Similarly, we apply the first-order Taylor expansion to convert \( \| q_i[t] - q_j[t] \|^2 \) in constraint (20c) to a linear function since it is a convex function with respect to \( q_i \) and \( q_j \). This can be expressed as follows:
\[ \| q_i[t] - q_j[t] \|^2 \geq 2(q_i[t] - 1 - q_j[t] - 1)^T \]
\[ (q_i[t] - q_j[t]) - \| q_i[t] - q_j[t] - 1 \|^2. \] (41)

Afterwards, we can denote the above equation as follows:
\[ G_0[t] - 1| q_i[t] - q_j[t] | \triangleq 2(q_i[t] - 1 - q_j[t] - 1)^T \]
\[ (q_i[t] - q_j[t]) - \| q_i[t] - q_j[t] - 1 \|^2. \] (42)

**Algorithm 1: SCA Algorithm for ARIS Deployment.**

**Input:** Initial feasible points \( \{ q_i^0, a_i^0, b_i^0 \}, r_k^\text{min}[t], d_{\text{min}}, \) iteration index \( i = 0, i_{\text{max}}, \) stopping criterion \( \varepsilon_1. \)

1. **repeat**
2. Set \( i \leftarrow i + 1. \)
3. Update \( q_i^i, a_i^i, b_i^i \) with given \( q_i^{i-1}, a_i^{i-1}, b_i^{i-1}. \)
4. Acquire \( \hat{E}(\hat{p}, \hat{p}') = \hat{h}(\hat{p}) - \hat{t}(\hat{p}, \hat{p}') \) based on (30).
5. Solve (43) to obtain \( \hat{p}'. \)
6. **until** \( |\hat{E}(\hat{p}') - \hat{E}(\hat{p}'')| \leq \varepsilon_1 \) or \( i > i_{\text{max}}. \)

**Output:** Optimal ARIS deployment \( q^* \).

Finally, we can substitute (42) into (20c), and problem (23) can be rewritten as follows:
\[ \min_{q,a,b,p} -\hat{E}(\hat{p}, \hat{p}') \] (43a)

\[ \text{s. t.} \]
\[ \hat{r}[t] + H_{0ab}[t]H'[t]H_{0ab}[t] \]
\[ - 2R \left[ H_{0ab}[t]H'[t]H_{0ab}[t] \right] \leq 0, \forall t \in T \] (43b)
\[ G_0[t - 1]| q_i[t] - q_j[t] | \geq d_{\text{min}}, \forall i,j \in N, i \neq j, \forall t \in T \] (43c)
\[ (20b), (39), (40). \] (43d)

Problem (43) becomes a convex optimization problem, which we solve by using CVXPY solver in python programming. The overall algorithm of the SCA method is shown in Algorithm 1.

**B. Joint ARIS Reflecting Elements On/off States and Phase Shift Problem**

For the given ARIS deployment \( q \) and power control \( g \), the sub-problem \( P^2 \) can be represented as follows:
\[ P^2 : \max_{\Delta, t} E(\Delta, t) \] (44a)

\[ \text{s. t.} \]
\[ r_k[t] \geq r_k^\text{min}[t], \forall k \in K, \forall n \in N', \forall i \in T_n, \forall t \in T \] (44b)
\[ 0 \leq t_n[i] < 2\pi, \forall n \in N', \forall i \in T_n, \forall t \in T \] (44c)
\[ \delta_{n}[i] \in \{ 0, 1 \}, \forall n \in N', \forall i \in T_n, \forall t \in T \] (44d)

This problem is still mixed-integer, non-convex, and quite challenging to solve in polynomial time since the information of the environment is unknown. Moreover, the real-time ARIS reflecting elements on/off states requires extensive computation and hardware cost, and conventional optimization methods cannot be applied. The exhaustive search method can be used to find the optimal solution; however, it is impractical for large-scale networks. Due to these reasons, we propose a DRL approach to solve sub-problem \( P^2 \). The reason we do not apply DRL for the whole optimization problem is that the action spaces combined for all ARIS deployment, ARIS reflecting elements on/off states, phase shift, and power control matrices will be too large and demands high computational cost. Here, we implement Actor-Critic Proximal Policy Optimization (AC-PPO) [49] as
it always provides an improved policy by using data that are currently accessible by the agent and thereby ensuring data efficiency and reliable performance. It could also be utilised in the environments where action spaces are discrete or continuous. Typically, since DRL is interpreted as Markov Decision Process (MDP), we first need to define state space \( S \), action space \( A \), and reward \( R \).

1) State Space: For each state at time \( t \), \( s_t \in S \) can be expressed as the tuples of the users’ locations and ARISs’ locations, the channel gain of the direct link, the channel gain of the ARIS-UE link and BS-ARIS link, and power control at time \( t \), respectively, and can be denoted as \( s_t = \{ q_k[t], q_n[t], h_{B,k}[t], h_{n,k}[t], H_{B,n}[t], g_k[t], \forall k \in K, \forall n \in N, \forall i \in I_n \} \).

2) Action Space: The action at time \( t \), \( a_t \in A \) contains the combination of the ARIS reflecting elements on/off states vector \( \delta_{in}[t] \), and phase shift vector \( \theta_{in}[t] \) at time \( t \), and can be denoted as \( a_t = \delta_{in}[t], \theta_{in}[t], \forall n \in N, \forall i \in I_n \} \).

3) Reward: Since the goal of our system is to maximize the energy efficiency, our reward function is defined as

\[
\tilde{R}_t(s_t|a_t) = \begin{cases} -1, & \text{if } \sum_{k=1}^{K} r_k[t] < r_k^{\text{min}}[t], \\ \mathcal{C}(\Delta, \Theta), & \text{otherwise}. \end{cases} \tag{45}
\]

As shown in Fig. 3, in our AC-PPO algorithm, the states information, \( s_t \) from the environment is obtained by the agent at the BS, and the agent observes and monitors the status of the location of the users and ARISs, channel gain of the links, and power control for each user. The agent includes the actor model and the critic model [50]. The actor model has the stochastic policy model \( \pi_{\psi}(a_t|s_t) \) with its own parameter \( \psi \) and learns to take action under the observation of the input states. The policy \( \pi_{\psi}(a_t|s_t) \) takes the observed states \( s_t \) from the environment as an input and suggests actions \( a_t \) to take as an output, and calculates the immediate reward \( \tilde{R}_t(s_t|a_t) \) depending on the action taken. The reward then provides as feedback to the agent, and the new state information \( s(t + 1) \) is obtained. Taking into account of the requirements for the users, under the given policy \( \pi_{\psi}(a_t|s_t) \) and reward function \( \tilde{R}_t(s_t|a_t) \), the cumulative discounted reward function at time \( t \) can be denoted as follows:

\[
V^{\pi_{\psi}}(s_t) = \mathbb{E}_t \left[ \sum_{t'=t}^{T-1} \xi^{t'-t} \tilde{R}_{t'}(s_{t'}|a_{t'}) \right], \forall s_t \in S, \tag{46}
\]

where \( 0 < \xi < 1 \) is the discount factor to prevent the total reward from reaching to infinity.

Moreover, the critic model contains the advantage function, \( \tilde{A}_t \), which is the estimate of the relative value of the selected action in the current state is defined as [51]:

\[
\tilde{A}_t = V^{\pi_{\psi}}(s_t) - b(s_t), \forall s \in S, \tag{47}
\]

where \( b(s_t) \) is the baseline estimate value function which provides the estimate of the discounted return starting from the current state \( s_t \).

The surrogate objective function of AC-PPO is to find the policy that maximizes the total rewards from the environment and can be expressed as follows [49]:

\[
L^{\text{CLIP}}(\psi) = \mathbb{E}_t \left[ \min \left( r_t(\psi) \tilde{A}_t, \text{clip}(r_t(\psi), 1 - \epsilon, 1 + \epsilon) \tilde{A}_t \right) \right], \tag{48}
\]

where

\[
r_t(\psi) = \frac{\pi_{\psi}(a_t|s_t)}{\pi_{\psi_{old}}(a_t|s_t)}.
\]

means the probability ratio. Given the states and actions, \( r_t(\psi) > 1 \) is the action is more plausible currently than it was in the old version of the policy, and \( 0 < r_t(\psi) < 1 \) if it is less plausible, and \( \epsilon \) is the clipping parameter. The clipping part of the objective function ensures that the PPO does not always favor actions with positive advantage and/or consistently avoid actions with negative advantage. The overall algorithm of the AC-PPO algorithm is described in Algorithm 2.

C. Power Control Problem

For the fixed ARIS deployment \( q \), ARIS reflecting elements on/off states \( \Delta \), and phase shift \( \Theta \), sub-problem \( P3 \) can be represented as follows:

\[
P3: \max_{g} \mathcal{E}(g) \tag{49a}
\]
Algorithm 2: AC-PPO Algorithm for ARIS Reflecting Elements on/off States and Phase Shift.

**Input:** Network states \( s_t \), learning rate, discount factor \( \xi \), clipping parameter \( \epsilon \).

1. **Initialization** Base policy \( \pi_\psi (a_t | s_t) \) with random parameters \( \psi \) and clipping parameter \( \epsilon \) and initial value function \( V^{\pi_0} (s_t) \)

2. **for** \( k \in K \) **do**
3. **for** each episode \( t \in T \) **do**
4. Collect the network observations: ARIS deployments \( q \) from Algorithm 1 and power control \( g \) from Algorithm 3 to achieve the initial state \( s_0 \)
5. **for** each \( t \in T \) **do**
6. Forward the network states \( s_t \in S \) to the AC-PPO algorithm
7. Observe the input states \( s_t \) and run the actor network
8. Select action \( a_t \in A \) based on policy \( \pi_\psi (a_t | s_t) \)
9. Obtain the reward \( R_t (s_t | a_t) \) and \( s_{t+1} \)
10. Calculate the probability ratio, \( r_t \)
11. Compute \( A_t \) based on current \( V^{\pi_0} (s_t) \) at the critic network according to (47)
12. Compute \( LCLIP (\psi) \) according to (48)
13. Update \( \pi_{\psi_{old}} \leftarrow \pi_\psi \)
14. **end for**
15. **end for**
16. **end for**

**Output:** Optimal AC-PPO network with \( \Delta^*, \Theta^* \).

\[
\begin{align*}
\text{s.t. } r_k [t] & \geq r_k^{\min}[t], \forall k \in K, \forall n \in N, \forall i \in I_n, \forall t \in T, \\
\text{tr}(g[t]^H g[t]) & \leq P_{\max}, \forall t \in T. 
\end{align*}
\] (49b)

(49c)

Sub-problem P3 is still a non-convex and NP-hard problem due to constraint (49b). Therefore, it is challenging to obtain the solutions in polynomial time. Therefore, we adopt the Whale Optimization Algorithm (WOA) to solve sub-problem P3. The WOA is a meta-heuristic algorithm that mimics the whales’ hunting strategy. The WOA has substantial advantages. First, unlike gradient-based algorithms, which involve computing and updating the gradients and step size throughout every iteration of the optimization process, WOA allows for such computation to be relaxed. Second, WOA is not influenced by the initial feasible solutions, which might have a significant impact on the convergence. Therefore, it has recently gained popularity among the research community due to it being an efficient optimizer.

The WOA algorithm includes two states: 1) the exploitation state (the encircling prey method and spiral bubble-net attacking method), and 2) the exploration state (the searching prey method). The detailed explanation of each state can be further described in the following subsections [52], [53], [54].

1) Exploitation State: The exploitation state of WOA includes two fundamental methods: the encircling prey method, and the spiral bubble-net attack method, which are discussed as follows:

**Encircling Prey Method:** Once the whales detect the location of their preys, they encircle them. Theoretically, the location of the prey is unknown in the search space, therefore, WOA assumes that the current best search agent is the target prey (optimum or close to optimum). The other whales (search agents) update their locations towards the best search agent. This behaviour can be mathematically implemented as follows [52]:

\[
\begin{align*}
\hat{D} &= \left| \hat{C} \cdot \hat{g}(\hat{j}) - \hat{g}(\hat{j}) \right|, \\
\hat{g}(\hat{j} + 1) &= \hat{g}(\hat{j}) - \hat{A} \cdot \hat{D},
\end{align*}
\] (50)

(51)

where \( \hat{g} \) is the location of the best search agent, \( \hat{j} \) is the current iteration, \( | \cdot | \) is the absolute value, \( \hat{C} \) and \( \hat{A} \) are coefficient vectors, and are computed as follows:

\[
\begin{align*}
\hat{A} &= 2 \hat{a} \cdot \hat{r} - \hat{a}, \\
\hat{C} &= 2 \cdot \hat{r},
\end{align*}
\] (52)

(53)

where \( \hat{r} \) is the random vector between 0 to 1, and \( \hat{a} \) is the control parameter vector linearly declining from 2 to 0 over the iterations, both in exploitation and exploration states. The aim of (52) and (53) is to balance between exploitation and exploration. When \( A \geq 1 \), WOA will perform exploration, and exploitation is done when \( A < 1 \).

**Spiral Bubble-net Attack Method:** This method combines both the shrinking encircling mechanism and the spiral movement mechanism of whales. Its purpose is to update the new location to fall between the current agent’s location and the best search agent. To mimic the helical shape movement of the whales, the equation can be expressed as

\[
\begin{align*}
\hat{D}' &= \left| \hat{g}'(\hat{j}) - \hat{g}(\hat{j}) \right|, \\
\hat{g}(\hat{j} + 1) &= \hat{D}' \cdot e^{bj} \cdot \cos(2\pi l) + \hat{g}(\hat{j}),
\end{align*}
\] (54)

(55)

where \( \hat{D}' \) indicates the distance between the current search agent and the target prey. Moreover, \( b \) is the constant for defining the shape of the logarithmic spiral, and \( l \) is the random number between \(-1 \) and \( 1 \). Here, coefficient vector \( \hat{A} \) is updated by setting random values in \([-1, 1] \).

Conventionally, once the whales locate the prey, they approach it using either the shrinking encircling method or the spiral bubble-net method synchronously. To imitate this synchronous behaviour, we set the 50% probability of choosing between these two methods to update the location of the whales for the optimization. Mathematically, it can be modeled as follows:

\[
\hat{g}(\hat{j} + 1) = \begin{cases} 
\hat{g}(\hat{j}) - \hat{A} \cdot \hat{D}, & \text{if } p < 0.5, \\
\hat{D}' \cdot e^{bj} \cdot \cos(2\pi l) + \hat{g}(\hat{j}), & \text{if } p \geq 0.5,
\end{cases}
\] (56)

where \( p = [0, 1] \) is the random number to represent the probability to choose between two mechanisms. When \( p < 0.5 \), WOA chooses the shrinking encircling mechanism, and if \( p \geq 0.5 \), WOA chooses the spiral movement mechanism.

2) Exploration State: The exploration state of WOA includes the searching for prey method. This state is necessary to prevent
Algorithm 3: WOA for Power Control

**Input:** Current power control \( g \), given \( q, \Delta, \) and \( \Theta \);

1. **Initialization** At iteration \( j = 1 \), initialize the total number of whale population \( g_u \), where \( u = \{1, \ldots, U\} \), and maximum number of iteration \( J_{max} \).
2. According to (59), calculate the fitness of the search agents \( g_u \), and identify the best search agent \( g^* (0) \).
3. **repeat**
   4. for \( u \leftarrow 1 \) to \( U \) (the number of whales) do
      5. Update \( a, A, C, l \) and \( p \).
      6. if \( p < 0.5 \) then
         7. Update \( \tilde{D} \) by (50) and \( \tilde{g} \) by (51).
      else
         8. Select a random \( \tilde{g}_{rand} \) and update \( \tilde{D} \) by (57).
         9. Update the location \( \tilde{g} \) by (58).
      end if
   10. end if
   11. Update \( \tilde{D} \) by (54) and \( \tilde{g} \) by (55).
12. **end if**
13. **else**
14. **end for**
15. Calculate the fitness of each search agent by (59).
16. Update the location of the best search agent \( \tilde{g}^* (j) \).
17. Update \( j \leftarrow j + 1 \).
18. **until** \( j > J_{max} \)

**Output:** Optimal power control \( g^* \).

the solution from being trapped at the local optimum, and failing to achieve the global optimum.

**Searching for Prey Method:** This approach is similar to encircling prey method in the exploitation state, but instead of claiming the location of the best search agent, and here, a random location is selected to update the locations of other search agents. It can be mathematically represented as follows:

\[
\tilde{D} = \left| \tilde{C} \cdot \tilde{g}_{rand}(\tilde{j}) - \tilde{g}(\tilde{j}) \right| , \tag{57}
\]

\[
\tilde{g}(\tilde{j} + 1) = \tilde{g}_{rand}(\tilde{j}) - \tilde{A} \cdot \tilde{D}, \tag{58}
\]

where \( \tilde{g}_{rand}(\tilde{j}) \) is the location of the search agent randomly selected from the search space.

Since the WOA algorithm is designed only for unconstrained optimization, we apply the penalty method to our sub-problem \( P3 \) in order to deal with the minimum achievable date rate constraint (49b) in the problem [54]. In our scenario, UEs are considered as search agents, and the power control of the BS \( g \) represents the location of the search agents. At each iteration \( \tilde{j} \), the power control \( g \) can be updated by either the encircling prey method, spiral bubble-net attack method, or searching for prey method. The fitness function which evaluates the optimal search agent can be expressed as follows:

\[
\text{Fitness}(g) = -\frac{R(g)}{P(g)} + \varepsilon \sum_{k=1}^{K} F_k(f_k(g))f_k^2(g), \tag{59}
\]

Algorithm 4: Proposed Joint ARIS Deployment, ARIS Reflecting Elements on/off States, Phase Shift, Power Control Optimization Algorithm.

1. **Initialization:** At \( \tau = 0 \), initialize \( \chi = 10^{-4} \) and variables \( q(0), \Delta (0), \Theta (0), g(0) \);
2. **repeat**
3. By applying Algorithm 1, solve problem \( P1 \) for given \( \Delta(\tau), \Theta(\tau), g(\tau) \) to obtain \( q(\tau + 1) \).
4. By applying Algorithm 2, solve problem \( P2 \) for given \( q(\tau + 1), g(\tau) \) to obtain \( \Delta(\tau + 1), \Theta(\tau + 1) \).
5. By applying Algorithm 3, solve problem \( P3 \) for given \( q(\tau + 1), \Delta(\tau + 1), \Theta(\tau + 1) \) to obtain \( g(\tau + 1) \).
6. Update \( \tau \leftarrow \tau + 1 \).
7. **until** \( |E(q, \Delta, \Theta, g)(\tau + 1) - E(q, \Delta, \Theta, g)(\tau)| \leq \chi \)
   Then, set \( q(\tau + 1), \Delta(\tau + 1), \Theta(\tau + 1), g(\tau + 1) \) as the desired solution.

where \( f_k(g) = r_k^{min}[t] - r_k[t] \) is the inequality function, and \( \varepsilon \) is the penalty factor coefficient. Since our sub-problem \( P3 \) is the maximization problem, we add the negative sign ahead of the objective function to convert it into a minimization problem. The index function \( F_k(f_k(g)) = 1 \) if \( f_k(g) < 0 \), and \( F_k(f_k(g)) = 0 \) if \( f_k(g) \geq 0 \). The pseudo-code of our WOA based power control can be described as in Algorithm 3.

D. Overall Algorithm Complexity Analysis

The overall iterative algorithm for solving our optimization problem (18) is described in Algorithm 4 with the aforementioned proposed solutions to three sub-problems. According to the results in [31], [48] and [54], the complexity of our solutions can be obtained by each algorithm for each sub-problem. For the ARIS deployment sub-problem, the SCA is adopted as in Algorithm 1. Since there are \( K \) users, the computational complexity of the SCA method is obtained as \( O_1(K^{3.5}\log(1/\varepsilon_1)) \) where \( \varepsilon_1 \) is the variable to control the accuracy of the SCA algorithm. For the AC-PPO algorithm for ARIS reflecting elements on/off states and phase shift as in Algorithm 2, the computational complexity is \( O_2(\alpha^2K) \), where \( \alpha \in A \) is the total number of actions taken by the agent. With WOA for power control as in Algorithm 3, the computational complexity is \( O_3(JU(m + K)) \), where \( J \) is the number of iterations for WOA, \( U = 30 \) denotes the number of whale populations, and \( m \) represents the number of inequality constraints in sub-problem \( P3 \). Henceforth, the overall computational complexity for solving (18) can be acquired as \( O(\hat{\tau}K^{3.5}\log(1/\varepsilon_1) + \hat{\tau}\alpha^2K + \hat{\tau}JU(m + K)) \), where \( \hat{\tau} \) denotes the number of iterations for Algorithm 4.

V. PERFORMANCE EVALUATION

In this section, we evaluate our proposed technique of energy-efficient multiple ARISs-assisted downlink communication system via numerical analysis. The network design comprises of 12 UEs uniformly distributed within 100 m x 100 m square region and the BS with 15 multiple antennas located at the center of the
TABLE II
SIMULATION PARAMETERS

| Parameter                                      | Value |
|------------------------------------------------|-------|
| Number of reflecting elements on ARIS $n_e$ | 10    |
| Bandwidth $W$                                 | 10 MHz [55] |
| Noise power $\sigma^2$                        | -174 dBm |
| Path loss exponent $\alpha$                  | 4     |
| Channel gain at reference distance $\kappa$  | -40 dBm |
| Rician factor $R$                             | 10    |
| Circuit power of each AF relay in UAV $P_{UI}$| 10 dBm [41] |
| BS power amplifier efficiency $\mu$           | 0.8 [41] |
| Circuit power of each user $P_{ui}^{max}$     | 10 dBm [41] |
| Circuit power of each RIS element $P_{RIS}$   | 10 dBm [41] |
| Circuit power of BS $P_B$                     | 39 dBm |
| Coefficient of profile drag $\nu$             | 0.012 [56] |
| Density of air $\rho$                         | 1.225 kg/m$^3$ [56] |
| Rotor solidity $\Lambda$                      | 0.05 [56] |
| Disc area $\eta$                              | 0.3 m$^2$ [56] |
| Blade angular velocity $\omega$               | 300 rad/s [56] |
| Radius of rotor $\rho$                        | 0.4 [56] |
| Incremental correction factor $\epsilon$      | 0.05 |
| UAV weight                                     | 20 kg [57] |
| Speed of UAV                                   | 18 m/s [44] |
| Clipping parameter $\epsilon$                 | 0.2 |
| Learning rate                                  | 0.0002 |
| Discount factor $\xi$                         | 0.9 |
| Mini batch size                                | 64 |
| Number of episodes                             | 1,000 |
| Number of time steps                           | 300,000 |

coverage area. There are 4 ARISs to support communication, and each ARIS is integrated with 10 reflecting elements. The ARISs can hover at a maximum altitude of 100 m. The simulation parameters can be observed in Table II.

To evaluate our proposed algorithm, we compare our method with four benchmark schemes, which are explained as follows:

- **Single-ARIS:** In this scheme, we implement a single ARIS instead of using multiple ARISs to support the downlink communications from BS. Specifically, the number of reflecting elements in the Single-ARIS case is assigned to be the same as the total number of reflecting elements for all multiple ARISs in the proposed scheme. The optimization problem is then solved by using our proposed algorithms.

- **ARIS (NPS):** In this approach, we deploy 4 ARISs with the fixed phase shifts. The ARIS deployment problem is solved by SCA, the ARIS reflecting elements on/off states problem is solved by AC-PPO, and transmit power allocation problem is solved by WOA alternatingly.

- **Random:** In this design, the deployment of ARISs is random. The optimal phase shift is optimized by the AC-PPO algorithm without the optimal reflecting elements on/off states, and optimal transmit power of the BS.

- **UAV-Relay:** In this method, ARIS is not used. Instead, 4 UAVs are deployed as amplify-and-forward relays and the incident signal is linearly processed, and re-transmit them toward the required destination [58]. Under the UAV-Relay scheme, we consider that $N$ UAV-relays equipped with $I$ antennas are deployed in the same approach as ARISs [41]. The optimization problem is then solved by using our proposed algorithms.

Fig. 4 compares the average sum-rate of users with our proposed DRL-based algorithm towards the above-mentioned benchmark schemes. In all circumstances, the average sum-rate increases as the maximum transmit power rises. Our proposed algorithm outperforms ARIS (NPS) by 18% and single-ARIS by 51%, respectively. This demonstrates how the multiple ARISs can achieve better outcomes than a single ARIS since it can provide several paths between the BS and UEs. Our algorithm outperforms most benchmark schemes in average sum-rate except for the UAV-relay scenario. UAV-relay provides the highest performance since it processes and re-transmits the incident signal using a dedicated power source. As a consequence, it consumes more energy which can be observed in Fig. 5.

Fig. 5 depicts the comparison of the energy efficiency under different algorithms. The smooth data is demonstrated by the solid curved line, which represents the Savitzky-Golay filter. In all scenarios, energy efficiency increases faster until the transmit power of the BS reaches to 10 dBm. Since then, the energy efficiency hasn’t improved much as the function does not increase monotonically with respect to the transmit power. Our proposed algorithm achieves 68% increase compared to the UAV-relay scenario and 30% increase compared to the single-ARIS scenario.
Next, we evaluate the convergence of our proposed AC-PPO algorithm with different values of $P_{\text{max}}$ ranging from 0 dBm to 40 dBm. As shown in Fig. 6, it can be observed that in all cases, the convergence of our cumulative rewards increases with respect to the increase in transmit power. We can see a significant difference in the performance when $P_{\text{max}}$ is low, and the performance difference becomes lesser as $P_{\text{max}}$ becomes higher. This suggests that SINR has a significant impact on the overall performance of the cumulative rewards.

Following that, we examine how various learning rates affect on our cumulative rewards, ranging from the set of {0.02, 0.002, 0.0002, 0.00002}. As seen in Fig. 7, a higher learning rate does not enable our cumulative reward to converge faster and provides less performance. Although it takes longer to converge, the learning rate of 0.0002 delivers better performance than 0.002 and 0.0002. In this case, we chose a learning rate of 0.0002 since it produces the highest cumulative rewards for our proposed method.

Furthermore, we compare the spectral efficiency of our proposed multiple ARISs-assisted system to that of multiple RISs-assisted systems. In this approach, we employ 4 RISs on the ground level rather than mounted on the UAVs [59]. Fig. 8 demonstrates the cumulative distribution function (CDF) values of average sum-rates. As seen in Fig. 8, our proposed system achieved 33% performance increase compared to RISs-assisted system. This is because our proposed system takes into account the deployment of UAVs, which provides improved LOS communications between the BS and UEs. Concurrently, we experiment the performance of spectral efficiency with different numbers of reflecting elements. In all scenarios, the results show that as the number of reflecting elements increases, so do the UEs’ average sum-rates. This indicates that the spectral efficiency will be improved by increasing the number of reflecting parts.

Moreover, we examine the energy efficiency with various ARIS numbers and different numbers of UEs, respectively. As shown in Fig. 9, we can observe that the energy efficiency improves as the number of ARIS components increases. Moreover, when the number is low, the energy efficiency improvement is significantly more compared to a larger number of ARIS. This indicates that for the small cell network with 12 UEs, we do not need to install a large amount of ARISs. Next, as shown in Fig. 10, we can observe that the energy efficiency almost linearly increases with the increasing number of UEs between 6 to 12.
energy efficiency, we formulated a joint ARIS deployment, ARIS reflecting elements on/off states, phase shift, and power control problem. As the formulated problem is MINLP and NP-hard, we decompose our problem into three sub-problems: ARIS deployment problem, joint reflecting elements on/off states and phase shift problem, and power control problem. We then proposed the SCA approach, the AC-PPO method and the WOA to solve our sub-problems alternatingly. Through extensive numerical analysis, we have proved that by integrating multiple ARISs in the downlink communication system, it can significantly outperform several benchmark schemes; especially in spectral efficiency compared to a multiple RISs-assisted scenario and energy efficiency compared to a single ARIS-assisted scenario.

VI. CONCLUSION

In this article, we have studied an energy-efficient multiple ARISs-assisted downlink communication system. To maximize

---

Fig. 10. CDF of energy efficiency with different number of UEs.

Fig. 11. Comparison of energy-efficiency with different on/off states.

We can perceive that more ARISs and more UEs help improve the energy efficiency of the multiple ARISs-assisted downlink communication system.

Finally, we investigate the impact of the on/off states of the reflecting elements on energy efficiency. We consider the following two scenarios: 1) without considering the on/off states of reflecting elements (Wo on/off), i.e., $\Delta_{n}$ is an identity matrix for each ARIS $n$, and 2) uniform random on/off states (Random on/off). The remaining optimization problems are solved with our proposed algorithms. As illustrated in Fig. 11, in all scenarios, the energy efficiency improves with the increase in ARIS elements since, with more ARIS elements, the system can create more focused and directed beams to the users, reducing the energy consumption required to transmit a signal over long distances. Among them, our proposed system exceeds the others because, in the first case, the reflecting elements remain active during the entire transmission period. For the random on/off case, it does not consider the optimal switching strategy. Our simulation results demonstrate an improvement in energy efficiency, which confirms the effectiveness of our proposed joint optimization system.

REFERENCES

[1] U. Cisco, “Cisco annual internet report (2018–2023) white paper,” Mar. 2020.
[2] Y. K. Tun, Y. M. Park, N. H. Tran, W. Saad, S. R. Pandey, and C. S. Hong, “Energy-efficient resource management in UAV-assisted mobile edge computing,” IEEE Commun. Lett., vol. 25, no. 1, pp. 249–253, Jan. 2021.
[3] X. Hu, K.-K. Wong, K. Yang, and Z. Zheng, “UAV-assisted relaying and edge computing: Scheduling and trajectory optimization,” IEEE Trans. Wireless Commun., vol. 18, no. 10, pp. 4738–4752, Oct. 2019.
[4] M. A. ElMossallamy, H. Zhang, L. Song, K. G. Seddik, Z. Han, and G. Y. Li, “Reconfigurable intelligent surfaces for wireless communications: Principles, challenges, and opportunities,” IEEE Trans. Cogn. Commun. Netw., vol. 6, no. 3, pp. 990–1002, Sep. 2020.
[5] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, “Reconfigurable intelligent surfaces for energy efficiency in wireless communication,” IEEE Trans. Wireless Commun., vol. 18, no. 8, pp. 4157–4170, Aug. 2019.
[6] Y. Liu et al., “Reconfigurable intelligent surfaces: Principles and opportunities,” IEEE Commun. Surv. Tuts., vol. 23, no. 3, pp. 1546–1577, thirdquarter 2021.
[7] L. Li et al., “Machine-learning reprogrammable metasurface imager,” Nature Commun., vol. 10, no. 1, pp. 1–8, Mar. 2019.
[8] M.-H. T. Nguyen, E. Garcia-Palacios, T. Do-Duy, O. A. Dobre, and T.-Q. Duong, “UAV-aided aerial reconfigurable intelligent surface communications with massive MIMO system,” IEEE Trans. Cogn. Commun. Netw., vol. 8, no. 4, pp. 1828–1838, Dec. 2022.
[9] B. Shang, R. Shafin, and L. Liu, “UAV swarm-enabled aerial reconfigurable intelligent surface (SARIS),” IEEE Wireless Commun., vol. 28, no. 5, pp. 156–163, Oct. 2021.
[10] M. Li, X. Tao, N. Li, and H. Wu, “Energy-efficient covert communication with the aid of aerial reconfigurable intelligent surface,” IEEE Commun. Lett., vol. 26, no. 9, pp. 2101–2105, Sep. 2022.
[11] T. Jiang, L. Niu, X. Tang, X. Tang, and D. Zhai, “Aerial reconfigurable intelligent surface-assisted secrecy: A learning approach,” IEEE Commun. Lett., vol. 26, no. 1, pp. 18–22, Jan. 2022.
[12] L. Ge, H. Zhang, J.-B. Wang, and G. Y. Li, “Reconfigurable wireless relaying with multi-UAV-carried intelligent reflecting surfaces,” IEEE Trans. Veh. Technol., vol. 72, no. 4, pp. 4932–4947, Apr. 2023.
[13] B. Shang, E. S. Bentley, and L. Liu, “UAV swarm-enabled aerial reconfigurable intelligent surface: Modeling, analysis, and optimization,” IEEE Trans. Commun., vol. 71, no. 6, pp. 3621–3636, Jun. 2023.
[14] I. Bacaille, S. Herbin, A. Munari, R. Hermenier, T. Rasheed, and S. Allsopp, “Rapidly deployable network for tactical applications: Aerial base station with opportunistic links for unattended and temporary events ABSOLUTE example,” in Proc. IEEE Mil. Commun. Conf., 2013, pp. 1116–1120.
[15] P. Zhan, K. Yu, and A. L. Swindlehurst, “Wireless relay communications with unmanned aerial vehicles: Performance and optimization,” IEEE Trans. Aerosp. Electron. Syst., vol. 47, no. 3, pp. 2068–2085, Jul. 2011.
[16] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, “Unmanned aerial vehicle with underlaid device-to-device communications: Performance and tradeoffs,” IEEE Trans. Wireless Commun., vol. 15, no. 6, pp. 3949–3963, Jun. 2016.
[17] Q. Feng, E. K. Tameh, A. R. Nix, and J. McGeehan, “WLCp2-06: Modelling the likelihood of line-of-sight for air-to-ground radio propagation in urban environments,” in Proc. IEEE Glob. Commun. Conf., 2006, pp. 1–5.

[18] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, “Wireless communication using unmanned aerial vehicles (UAVs): Optimal transport theory for hover time optimization,” IEEE Trans. Wireless Commun., vol. 16, no. 12, pp. 8052–8066, Dec. 2017.

[19] M. Mozaffari, A. T. Z. Kasgari, W. Saad, M. Bennis, and M. Debbah, “Beyond 5G with UAVs: Foundations of a 3D wireless cellular network,” IEEE Trans. Wireless Commun., vol. 18, no. 1, pp. 357–372, Jan. 2019.

[20] Z. Wang, L. Duan, and R. Zhang, “Adaptive deployment for UAV-aided communication networks,” IEEE Trans. Wireless Commun., vol. 18, no. 9, pp. 4531–4543, Sep. 2019.

[21] C. H. Liu, Z. Chen, J. Tang, J. Xu, and C. Piao, “Energy-efficient UAV control for effective and fair communication coverage: A deep reinforcement learning approach,” IEEE J. Sel. Areas Commun., vol. 36, no. 9, pp. 2059–2070, Sep. 2018.

[22] G. Lee, M. Jung, A. T. Z. Kasgari, W. Saad, and M. Bennis, “Deep reinforcement learning for energy-efficient networking with reconfigurable intelligent surfaces,” in Proc. IEEE Int. Conf. Commun., 2020, pp. 1–6.

[23] H. Zhang, B. Di, L. Song, and Z. Han, “Reconfigurable intelligent surfaces assisted communications with limited phase shifts: How many phase shifts are enough?,” IEEE Trans. Veh. Technol., vol. 69, no. 4, pp. 4498–4502, Apr. 2020.

[24] B. Di, H. Zhang, L. Song, Y. Li, Z. Han, and H. V. Poor, “Hybrid beamforming for reconfigurable intelligent surface based multi-user communications: Achievable rates with limited discrete phase shifts,” IEEE J. Sel. Areas Commun., vol. 38, no. 8, pp. 1809–1822, Aug. 2020.

[25] Y. Chen et al., “Reconfigurable intelligent surface assisted device-to-device communications,” IEEE Trans. Wireless Commun., vol. 20, no. 5, pp. 2792–2804, May 2021.

[26] Y. Ai, F. A. P. de Figueiredo, L. Kong, M. Cheffena, S. Chatzinotas, and B. Ottersten, “Secure vehicular communications through reconfigurable intelligent surfaces,” IEEE Trans. Veh. Technol., vol. 70, no. 7, pp. 7272–7276, Jul. 2021.

[27] A. Al-Hilo, M. Shokry, M. Elhattab, C. Assi, and S. Sharafeddine, “Reconfigurable intelligent surface enabled vehicular communication: Joint user scheduling and passive beamforming,” IEEE Trans. Veh. Technol., vol. 71, no. 3, pp. 2333–2345, Mar. 2022.

[28] C. Huang, R. Mo, and C. Yuen, “Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning,” IEEE J. Sel. Areas Commun., vol. 38, no. 8, pp. 1839–1850, Aug. 2020.

[29] H. Yang, Z. Xiong, J. Zhao, D. Niyato, L. Xiao, and Q. Wu, “Deep reinforcement learning for energy-efficient networking with reconfigurable intelligent surfaces,” IEEE Trans. Veh. Technol., vol. 21, no. 3, pp. 1806–1822, Mar. 2022.

[30] M. Samir, M. Elhattab, C. Assi, S. Sharafeddine, and A. Ghrayeb, “Optimizing age of information through aerial reconfigurable intelligent surfaces: A deep reinforcement learning approach,” IEEE Trans. Veh. Technol., vol. 70, no. 4, pp. 3978–3983, Apr. 2021.

[31] E. Khorov, A. Kiryanov, A. Lyakhov, and G. Bianchi, “A tutorial on IEEE 802.11 AX high efficiency WLANs,” IEEE Commun. Surv. Tuts., vol. 21, no. 1, pp. 197–216, Firstquarter 2019.

[32] E. Khorov, I. Levitsky, and I. F. Akyildiz, “Current status and directions of IEEE 802.11 be, the future Wi-Fi 7,” IEEE Access, vol. 8, pp. 88664–88688, 2020.

[33] Z. Yang et al., “Energy-efficient wireless communications with distributed reconfigurable intelligent surfaces,” IEEE Trans. Wireless Commun., vol. 21, no. 1, pp. 665–679, Jan. 2022.

[34] Z. Wang, L. Liu, and S. Cui, “Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis,” IEEE Trans. Wireless Commun., vol. 19, no. 10, pp. 6607–6620, Oct. 2020.

[35] C. Huang, G. C. Alexandropoulos, A. Zappone, M. Debbah, and C. Yuen, “Energy efficient multi-user MISO communication using low resolution large intelligent surfaces,” in Proc. IEEE Glob. Commun. Conf. Workshops, 2018, pp. 1–6.

[36] Y. Zeng, J. Xu, and R. Zhang, “Energy minimization for wireless communication with rotary-wing UAV,” IEEE Trans. Wireless Commun., vol. 18, no. 4, pp. 2339–2345, Apr. 2019.

[37] Q. Wu, Y. Zeng, and R. Zhang, “Joint trajectory and communication design for multi- UAV enabled wireless networks,” IEEE Trans. Wireless Commun., vol. 17, no. 3, pp. 2109–2121, Mar. 2018.

[38] Y. Wang, M. Chen, Z. Yang, T. Luo, and W. Saad, “Deep learning for optimal deployment of UAVs with visible light communications,” IEEE Trans. Wireless Commun., vol. 19, no. 11, pp. 7049–7063, Nov. 2020.

[39] G. Scutari and Y. Sun, “Parallel and distributed successive convex approximation methods for big-data optimization,” in Multiagent Optimization, 2022, Berlin, Germany: Springer, Nov. 2018, pp. 141–308.

[40] S. S. Hassan, Y. M. Park, Y. K. Tun, W. Saad, Z. Han, and C. S. Hong, “3TO: THz-enabled throughput and trajectory optimization of UAVs in 6G networks by proximal policy optimization deep reinforcement learning,” in Proc. IEEE Int. Conf. Commun., 2022, pp. 5712–5718.

[41] J. Schulman, P. Fusioli, P. Dhariwal, A. Radford, and O. Klimov, “Proximal policy optimization algorithms,” 2017, arXiv:1707.06347.

[42] H.-K. Lim, J.-B. Kim, J.-S. Heo, and Y.-H. Han, “Federated reinforcement learning for training control policies on multiple IoT devices,” Sensors, vol. 20, no. 5, Mar. 2020, Art. no. 1359.

[43] V. Mnih et al., “Asynchronous methods for deep reinforcement learning,” in Proc. Int. Conf. Mach. Learn., 2016, pp. 1928–1937.

[44] S. Mirjalili and A. Latifi, “Whale optimization algorithm,” Adv. Eng. Softw., vol. 95, pp. 51–67, May 2016.

[45] M. Mafarja and S. Mirjalili, “Whale optimization approaches for wrapper feature selection,” Appl. Soft Comput., vol. 62, pp. 441–453, Jan. 2018.

[46] Q.-V. Pham, S. Mirjalili, N. Kumar, M. Alazab, and W.-J. Hwang, “Whale optimization algorithm with applications to resource allocation in wireless networks,” IEEE Trans. Veh. Technol., vol. 69, no. 4, pp. 4285–4297, Apr. 2020.

[47] Z. Mohamed and S. Aissa, “Energy-efficient joint broadcast-unicast communications via dual-polarized aerial RIS,” IEEE Trans. Wireless Commun., vol. 22, no. 3, pp. 2113–2126, Mar. 2023.

[48] S. Jung, W. J. Yun, M. Shin, J. Kim, and J.-H. Kim, “Orchestrated scheduling and multi-agent deep reinforcement learning for cloud-assisted multi-UAV charging systems,” IEEE Trans. Veh. Technol., vol. 70, no. 6, pp. 5362–5377, Jun. 2021.

[49] Y. Zeng, R. Zhang, and T. J. Lim, “Wireless communications with unmanned aerial vehicles: Opportunities and challenges,” IEEE Commun. Mag., vol. 54, no. 5, pp. 36–42, May 2016.

[50] T. Zhang, G. Liu, H. Zhang, W. Kang, G. K. Karagiannidis, and A. Nallanathan, “Energy-efficient resource allocation and trajectory design for UAV relaying systems,” IEEE Trans. Commun., vol. 68, no. 10, pp. 6483–6498, Oct. 2020.

[51] P. S. Aung, Y. K. Tun, Z. Han, and C. S. Hong, “Energy-efficiency maximization of multiple RISs-enabled communication networks by deep reinforcement learning,” in Proc. ICC IEEE Int. Conf. Commun., 2022, pp. 2181–2186.
Yu Min Park (Graduate Student Member, IEEE) received the B.S. degree in applied mathematics and computer engineering and the M.S. degree in computer engineering in 2019 and 2021, respectively, from Kyung Hee University, Seoul, South Korea, where he is currently working toward the Ph.D. degree in computer engineering. His research interests include reinforcement learning, intelligent networking management system, and network resource optimization.

Zhu Han (Fellow, IEEE) received the B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 1997, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland, College Park, MD, USA, in 1999 and 2003, respectively. From 2000 to 2002, he was an R&D Engineer of JDSU, Germantown, MD. From 2003 to 2006, he was a Research Associate with the University of Maryland. From 2006 to 2008, he was an Assistant Professor with Boise State University, Boise, ID, USA. He is currently a John and Rebecca Moores Professor with the Electrical and Computer Engineering Department as well as the Computer Science Department, University of Houston, Houston, TX, USA. He has been a 1% highly cited Researcher since 2017 according to Web of Science. His main research targets on the novel game-theory related concepts critical to enabling efficient and distributive use of wireless networks with limited resources. His research interests include wireless resource allocation and management, wireless communications and networking, quantum computing, data science, smart grid, carbon neutralization, security and privacy. Dr. Han was the recipient of the NSF Career Award in 2010, Fred W. Ellersick Prize of the IEEE Communications Society in 2013, EURASIP Best Paper Award for the Journal on Advances in Signal Processing in 2015, IEEE Leonard G. Abraham Prize in the field of Communications Systems (best paper award in IEEE JSAC) in 2016, and several best paper awards in IEEE conferences. Dr. Han was an IEEE Communications Society Distinguished Lecturer from 2015 to 2018. He has been a AAAS Fellow since 2019, and ACM distinguished Member since 2019. Dr. Han is also the winner of the 2021 IEEE Kiyoshi Tomiyasu Award (an IEEE Field Award), for outstanding early to mid-career contributions to technologies holding the promise of innovative applications, with the following citation: for contributions to game theory and distributed management of autonomous communication networks.

Yan Kyaw Tun (Member, IEEE) received the B.E. degree in marine electrical systems and electronics engineering from Myanmar Maritime University, Thanlyin, Myanmar, in 2014, and the Ph.D. degree in computer science and engineering from Kyung Hee University, Seoul, South Korea, in 2021. From 2021 to 2023, he was a Postdoctoral Research Fellow with the Intelligent Networking Lab, Kyung Hee University, Seoul, South Korea, and with the Division of Network and Systems Engineering, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden. He is currently an Assistant Professor with the Department of Electronic Systems, Aalborg University, Aalborg, Denmark. His research interests include integrated network economics, game theory, optimization theory, machine learning, and next-generation wireless communication and networking. Dr. Tun was the recipient of the Best Ph.D. Thesis Award in Engineering in 2020 and best student paper award at the IEICE Asia-Pacific Network Operations and Management Symposium (APNOMS) in 2019.

Choong Seon Hong (Senior Member, IEEE) received the B.S. and M.S. degrees in electronic engineering from Kyung Hee University, Seoul, South Korea, in 1983 and 1985, respectively, and the Ph.D. degree from Keio University, Tokyo, Japan, in 1997. In 1988, he joined KT, Gyeonggi-do, South Korea, where he was involved in broadband networks as a Member of the Technical Staff. Since 1993, he has been with Keio University, Tokyo. He was with the Telecommunications Network Laboratory, KT, as a Senior Member of Technical Staff and as the Director of the Networking Research Team till 1999. Since 1999, he has been a Professor with the Department of Computer Science and Engineering, Kyung Hee University. His research interests include future Internet, intelligent edge computing, network management, and network security. Dr. Hong is a Member of the Association for Computing Machinery (ACM), Institute of Electronics, Information and Communication Engineers, Information Processing Society of Japan (IPSJ), Korean Institute of Information Scientists and Engineers, Korean Institute of Communications and Information Sciences (KICS), the Korean Information Processing Society (KIPS), and Open Standards and ICT Association (OSIA). He was the General Chair, the TPC Chair/Member, or an Organizing Committee Member of international conferences, such as the Network Operations and Management Symposium (NOMS), International Symposium on Integrated Network Management (IM), Asia-Pacific Network Operations and Management Symposium (APNOMS), End-to-End Monitoring Techniques and Services (E2EMON), IEEE Consumer Communications and Networking Conference (CCNC), Assurance in Distributed Systems and Networks (ADSN), International Conference on Parallel Processing (ICPP), Data Integration and Mining (DIM), World Conference on Information Security Applications (WISA), Broadband Convergence Network (BeN), Telecommunication Information Networking Architecture (TINA), International Symposium on Applications and the Internet (SAINT), and International Conference on Information Networking (ICOIN). He was an Associate Editor for IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT and IEEE JOURNAL OF COMMUNICATIONS AND NETWORKS. He was also the Guest Editor of IEEE NETWORK MAGAZINE. He is currently an Associate Editor for the International Journal of Network Management and Future Internet.