Active IRS Aided Multiple Access for Energy-Constrained IoT Systems

Guangji Chen, Qingqing Wu, *Senior Member, IEEE*, Chong He, *Member, IEEE*, Wen Chen, *Senior Member, IEEE*, Jie Tang, *Senior Member, IEEE*, and Shi Jin, *Senior Member, IEEE*

***Abstract*— In this paper, we investigate the fundamental multiple access (MA) scheme in an active intelligent reflecting surface (IRS) aided energy-constrained Internet-of-Things (IoT) system, where an active IRS is deployed to assist the uplink transmission from multiple IoT devices to an access point (AP). Our goal is to maximize the sum throughput by optimizing the IRS beamforming vectors across time and resource allocation. To this end, we first study two typical active IRS aided MA schemes, namely time division multiple access (TDMA) and non-orthogonal multiple access (NOMA), by analytically comparing their achievable sum throughput and proposing corresponding algorithms. Interestingly, we prove that given only one available IRS beamforming vector, the NOMA-based scheme generally achieves a larger throughput than the TDMA-based scheme, whereas the latter can potentially outperform the former if multiple IRS beamforming vectors are available to harness the favorable time selectivity of the IRS. To strike a flexible balance between the system performance and the associated signaling overhead incurred by more IRS beamforming vectors, we then propose a general hybrid TDMA-NOMA scheme with device grouping, where the devices in the same group transmit simultaneously via NOMA while devices in different groups occupy orthogonal time slots. By controlling the number of groups, the hybrid TDMA-NOMA scheme is applicable for any given number of IRS beamforming vectors available. Despite of the non-convexity of the considered optimization problem, we propose an efficient algorithm based on alternating optimization, where each subproblem is solved optimally. Simulation results illustrate the practical superiorities of the active IRS over the passive IRS in terms of the coverage extension and supporting multiple energy-limited devices, and demonstrate the effectiveness of our proposed hybrid MA scheme for flexibly balancing the performance-cost tradeoff.**

***Index Terms*— Active intelligent reflecting surface (IRS), multiple access (MA), IRS beamforming, resource allocation, throughput maximization.**

I. Introduction

With the rapid development of Internet-of-Things (IoT) technologies, the unprecedented proliferation of electronic tablets, wearable devices, and mobile sensors is set to continue, which hastens a variety of IoT applications such as smart transportation, smart metering, and smart cities [1]. In a typical IoT system, multiple sensor devices connect with an access point (AP) to form a wireless data collection system, which has been widely deployed in various practical applications, i.e., event detection for emergency services, external environment monitoring, wireless surveillance for public safety, etc., [2]. For these applications, the low-cost devices gather a large volume data from specific areas and then send it to the AP for further processing. However, the limited battery capacity of IoT devices is one of critical issues due to their practical size and cost constraints, which fundamentally limits their information uploading capabilities [3], [4].

To overcome the aforementioned limitations, new cost-effective wireless techniques have to be developed for assisting data transmission of IoT systems. Recently, intelligent reflecting surface (IRS) has emerged as a cost-effective technology for enhancing the spectral- and energy-efficiency of future wireless networks [5], [6], [7]. Generally, IRS technologies mainly involve two types of architectures, namely passive IRS and active IRS. In particular, passive IRS is a digitally-controlled meta-surface composed of a large number of low-cost passive elements in tuning the phase shifts of the incident signals. With the proper design of the phase shifts of each element, IRS is capable of enhancing the signal reception at the desired destinations and/or mitigating the interference to unintended users, thereby artificially establishing favorable propagation conditions without requiring any RF chains. The fundamental power scaling law of passive IRS was firstly unveiled in [8] and [9], which demonstrated that passive IRS can provide an asymptotic squared-power gain in terms of received power at users via passive beamforming. The above

Manuscript received 21 January 2022; revised 7 May 2022 and 18 July 2022; accepted 2 September 2022. Date of publication 20 September 2022; date of current version 10 March 2023. The work of Qingqing Wu was supported by FDCT under Grant 0119/2020/A3. The work of Wen Chen was supported in part by the National Key Project 2020YFB1807700 and Project 2018YFB1801102, in part by Shanghai Kewei under Grant 20JC1416502 and Grant 22JC1404000, and in part by NSFC under Grant 62071296. The associate editor coordinating the review of this article and approving it for publication was S. Prakriya. *(Corresponding author: Qingqing Wu.)*

Guangji Chen and Qingqing Wu are with the State Key Laboratory of Internet of Things for Smart City, University of Macau, Macau 999078, China (e-mail: guangjichen@um.edu.mo; qingqingwu@um.edu.mo).

Chong He and Wen Chen are with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: hechong@sjtu.edu.cn; wenchen@sjtu.edu.cn).

Jie Tang is with the School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510641, China (e-mail: eejtang@scut.edu.cn).

Shi Jin is with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: jinshi@seu.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TWC.2022.3206332.

Digital Object Identifier 10.1109/TWC.2022.3206332

advantages of passive IRS have then inspired an intensive research interest in optimizing IRS phase shifts for different wireless communication system setups, such as multi-cell cooperation [10], [11], [12], physical layer security [13], [14], [15], millimeter-wave communications [16], [17], and unmanned aerial vehicle communications [18], [19]. While these works aimed at exploiting passive IRS for enhancing wireless information transmissions (WIT) of cellular networks, its high passive beamforming gain is also practically appealing for multifarious IoT application scenarios and unlocking its full potential in extending the lifetime of energy-constrained IoT devices [20]. Regarding the IRS-enabled IoT systems, several works have been emerged on three typical research lines, namely IRS-aided wireless information and power transfer (SWIPT) [21], [22], [23], IRS-aided wireless powered communication networks (WPCNs) [24], [25], [26], [27], and IRS-aided mobile edge computing (MEC) [28], [29], [30].

Specifically, all of the aforementioned works have considered to exploit passive IRS for assisting wireless communications. Nevertheless, one critical issue of passive IRS-aided wireless systems is that its performance may be practically limited by the well-known high (product-distance) path-loss [6]. To address the issue of passive IRS, a new type of IRS, called active IRS, has been recently proposed in [31], [32], [33] and [34] by amplifying the incident signals with low-cost hardware. Different from passive IRS, active IRS comprises a number of active elements, which are equipped with low-cost negative resistance components (e.g., negative impedance converter and tunnel diode), thereby enabling to amplify the reflected signals [31]. It is worth noting that the active IRS is quite different from the amplify-and-forward (AF) relay although it is capable of amplifying incident signals. Specifically, for the conventional AF relay, power-consuming RF chains are generally needed to receive signals first and then transmit it with the power amplification. In contrast, the active IRS does not need RF chains and the basic operation mechanism of the active IRS is similar to that of the passive IRS, which directly amplify/reflect signals in the air with low-power reflection-type amplifiers. By leveraging the active IRS in wireless transmissions of cellular networks, the joint AP and IRS beamforming design problems were investigated in [31] and [32] for different system setups, i.e., single user uplink and multi-user downlink communication systems, respectively. In addition, the optimization of the active IRS deployment was studied in [33]. Benefited by smartly controlling the amplification gain at each element, the results in [31], [32] and [33] demonstrated that the active IRS can perform better than the passive IRS in most practical scenarios.

Despite the aforementioned advantages of applying active IRS for assisting cellular networks, the employment of active IRS in energy-constrained IoT systems is attractive for overcoming the issue of the low information uploading capabilities for IoT devices due to their limited battery capacity. Note that massive connectivity is another requirement of IoT systems. Regarding the issue of the massive access in IoT systems, multiple access (MA) technique is one of the most fundamental enablers for accommodating a large number of IoT devices. The existing MA techniques can be loosely classified into two categories, namely orthogonal multiple access (OMA) and non-orthogonal multiple access (NOMA). NOMA is practically appealing for IoT networks due to its capability to enable the access of massive devices by allowing multiple users to simultaneously access the same spectrum [35]. Compared to OMA, several superiorities have been shown by integrating NOMA into conventional IoT applications without active IRS, such as MEC [36], [37], [38] and data collection systems [39], [40].

Based on the above discussions, it is nature to investigate the potential performance gain of integrating active IRSs into energy-constrained IoT systems by considering different MA schemes. Regarding an active IRS-aided energy-constrained IoT system, several fundamental issues remain unsolved. First, does time division multiple access (TDMA) outperform NOMA in such systems? Since active IRS is able to proactively establish favorable time-varing wireless channels, it is generally believed that exploiting dedicated IRS beamforming patterns for each individual device has a beneficial effect for TDMA. Additionally, the amplification gain of each element at active IRS is highly dependent on the transmit power of devices. For the typical energy-limited scenario, the transmit power of each device when employing NOMA is generally lower than that of TDMA for a given amount of energy, which renders that NOMA may reap larger available amplification gains at the active IRS compared to TDMA. Taking the above factors into consideration, it still remains an open problem which MA scheme is more beneficial for maximizing the system throughput. Second, how to design a more advanced MA scheme that is capable of flexibly striking a balance between performance and signaling overhead? This question is driven by the fact that even if TDMA outperforms NOMA, it also incurs higher signalling overhead since more IRS beamforming patterns are needed. As such, it may not be preferable to rely on the pure TDMA-based scheme considering the performance-cost tradeoff, especially when the number of IoT devices is practically large.

Motivated by the above issues, we study an active IRS-aided energy-constrained IoT system considering different MA schemes, where an active IRS is deployed to assist the UL data transmission between multiple energy-constrained devices and an AP. Different from the conventional passive IRS, the ability of power amplification for the active IRS provides new degrees of freedom to combat against the severe double path loss and further enhance the received signal power. On the other hand, it also introduces new optimization variables and makes the IRS beamforming vectors and transmit power of each device closely coupled in the newly added IRS amplification power constraints, thus rendering the joint design of the IRS beamforming and resource allocation more challenging than that of the conventional passive IRS. The main contributions of this paper are summarized as follows.

- We first study active IRS-aided energy-constrained IoT systems by utilizing both TDMA and NOMA schemes. For the TDMA scheme, the IRS beamforming vectors can be adjusted dynamically across time for each individual device, whereas for the NOMA scheme, all the devices share the same set of IRS beamforming vector during

their data transmission. By utilizing the proposed models, we formulate the corresponding system sum throughput maximization problems by jointly optimizing the transmit power of each device, time allocation, and IRS beamforming vectors, subject to the energy constraints of IoT devices and the IRS amplification power constraints.

- For the TDMA scheme, we prove that the energy of each device would be used up for maximizing the sum throughput and then transform the original optimization problem into a more tractable one equivalently. An efficient algorithm is further proposed to solve it based on successive convex approximation (SCA), where all the variables are optimized simultaneously. For the NOMA scheme, we propose an alternating optimization (AO)-based method to partition the entire variables into two blocks, namely the transmit power of devices and IRS beamforming vectors. Based on semidefinite program (SDP) techniques and Charnes-Cooper transformations, each block of variables is obtained optimally in an iterative way until convergence is achieved.
- Regarding the achievable sum throughput of the active IRS aided TDMA and NOMA schemes, we prove that given only one available IRS beamforming vector, the NOMA based scheme generally achieves a larger throughput than the TDMA based scheme, whereas the latter can potentially outperform the former if multiple IRS beamforming vectors are available. To provide more flexibility for balancing the performance-cost tradeoff, we propose a hybrid TDMA-NOMA scheme, where multiple devices are partitioned into several groups and the devices in the same group transmit simultaneously via NOMA while devices in different groups occupy orthogonal time resources. The proposed scheme generalizes the TDMA and NOMA schemes as two special cases and is applicable for any given number of IRS beamforming vectors available. We further extend the AO-based method to solve its associated optimization problem by applying proper change of variables.
- Our numerical results validate the theoretical findings and demonstrate that the practical superiorities of the active IRS over the conventional passive IRS in terms of supporting multiple low-energy devices, extending coverage range, and reducing the required number of reflecting elements. Moreover, it is shown that our proposed hybrid TDMA-NOMA scheme is capable of significantly lowering the signaling overhead at the cost of slight performance loss by properly determining the number of devices in each group.

The rest of this paper is organized as follows. Section II presents the system model for the active IRS-aided energy-constrained IoT system and problem formulations considering TDMA and NOMA. Sections III introduces proposed efficient algorithms for the corresponding problems in Section II and provides the theoretical performance comparison for the TDMA and NOMA-based schemes. In Section IV, we propose a general hybrid TDMA-NOMA scheme and extend the AO-based algorithm for solving its associated optimization problem. Section V presents numerical results to evaluate the performance of our proposed schemes and draw useful insights. Finally, we conclude the paper in Section VI.

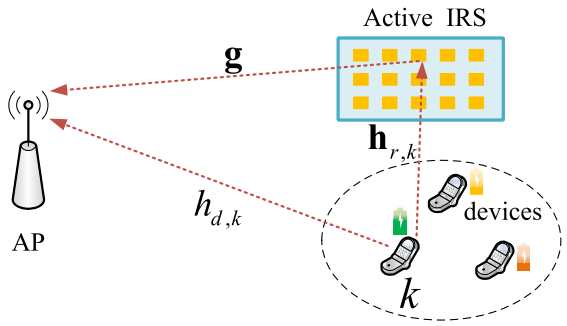


Fig. 1. An active IRS-aided energy-constrained uplink communication system.

Notations: Boldface upper-case and lower-case letter denote matrix and vector, respectively. $\mathbb{C}^{d_1 \times d_2}$ stands for the set of complex $d_1 \times d_2$ matrices. For a complex-valued vector $\mathbf{x}$, $\|\mathbf{x}\|$ represents the Euclidean norm of $\mathbf{x}$, $\arg(\mathbf{x})$ denotes the phase of $\mathbf{x}$, and $\mathrm{diag}(\mathbf{x})$ denotes a diagonal matrix whose main diagonal elements are extracted from vector $\mathbf{x}$. For a vector $\mathbf{x}$, $\mathbf{x}^*$ and $\mathbf{x}^H$ stand for its conjugate and conjugate transpose respectively. For a square matrix $\mathbf{X}$, $\mathrm{Tr}(\mathbf{X})$, $\|\mathbf{X}\|_2$ and $\mathrm{rank}(\mathbf{X})$ respectively stand for its trace, Euclidean norm and rank, while $\mathbf{X} \succeq \mathbf{0}$ indicates that matrix $\mathbf{X}$ is positive semidefinite. A circularly symmetric complex Gaussian random variable x with mean μ and variance σ^2 is denoted by $x \sim \mathcal{CN}(\mu, \sigma^2)$. $\mathcal{O}(\cdot)$ is the big-O computational complexity notation.

II. System Models and Problem Formulations

A. System Model

As shown in Fig. 1, we consider an active IRS-aided energy-constrained uplink communication system, where an active IRS with N elements, denoted by the set, $\mathcal{N} \triangleq \{1, \ldots, N\}$, is deployed to improve the efficiency of data transmission from K single-antenna IoT devices, denoted by the set $\mathcal{K} \triangleq \{1, \ldots, K\}$, to a single-antenna AP.[1] Without loss of generality, we focus on a typical energy-constrained IoT scenario, where a certain amount of energy E_k Joule (J) for device k is available at the beginning of each transmission period. For example, the energy sources for each device can be ambient energy sources (e.g., TV signals or a renewable energy source) or dedicated WPT. Then, all the devices use their available energy to transmit their own data to the AP in the uplink. In addition, all devices and the AP are assumed to operate over the same frequency band and the time duration of each transmission period is denoted by $T_{\max}$. Furthermore, let $\mathbf{g} \in \mathbb{C}^{N \times 1}$, $\mathbf{h}_{r,k} \in \mathbb{C}^{N \times 1}$, and $h_{d,k} \in \mathbb{C}$ denote the

[1]To unveil the potential benefits of the active IRS in energy-constrained uplink communication systems for improving the sum throughput, we assume that the AP is equipped with a single-antenna. Note that the AO principle in this paper is also applicable to the case of multiple antennas at the AP. Specifically, the optimal receive beamformers for the TDMA and NOMA schemes are maximum ratio combiner [8] and minimum mean square error-based form [36], respectively.

equivalent baseband channels from the IRS to the AP, from device k to the IRS, and from device k to the AP, respectively. All the wireless channels are assumed to be quasi-static flat-fading and thereby remain constant within each transmission period $T_{\max}$. To facilitate the fundamental performance comparisons for different MA schemes, we assume that the instantaneous CSI of all links is available by using channel estimation techniques proposed in [41]. Specifically, the AP first estimates the cascaded channels of devices-IRS-AP links based on the pilot signals sent by the devices. Based on the cascaded channel, the individual CSI of the device-IRS link and the IRS-AP link can be further reconstructed at the AP by using alternating least squares or vector approximate message passing methods [41]. Then, the algorithm of optimizing the IRS beamforming vectors and the related resource allocation is executed at the AP based on the obtained CSI. Finally, the AP sends the optimized IRS beamforming vectors and resource allocation results to the IRS controller and devices, respectively. Whether the active sensors are needed or not depends on the specific channel estimation technique. We have clarified that active sensors may not be needed with advanced channel estimation methods [41].

B. TDMA and NOMA-Based Multiple Access

For a typical transmission period, each device can transmit its own information signal to the AP. Furthermore, we propose two data transmission setups depending on whether TDMA or NOMA is used, as detailed below.

1) TDMA-Based Scheme: For the TDMA-based scheme, the AP receives information signals from different devices, which occupy orthogonal time slots (TSs). Let τ_k denote the time duration of the k-th TS, which is allocated for device k. Thus, we have $\sum_{k=1}^{K} \tau_k \leq T_{\max}$. In the k-th TS, a dedicated IRS beamforming pattern, denoted by $\mathbf{\Phi}_k = \mathrm{diag}\left(\phi_{k,1}, \ldots, \phi_{k,N}\right)$, is employed to reflect and amplify the transmitted signals. In particular, the reflecting/amplfication coefficient of the n-th element is denoted by $\phi_{k,n} = a_{k,n} e^{j\theta_{k,n}}, n \in \mathcal{N}$, where $a_{k,n}$ and $\theta_{k,n}$, $\theta_{k,n} \in [0, 2\pi)$, represent the corresponding amplitude and phase. It is worth noting that $a_{k,n}$ can be greater than 1 with active loads [31]. We denote the transmit power and the information bearing signal of device k as p_k and s_k, respectively, which satisfies $\mathrm{E}\left\{|s_k|^2\right\} = 1$. In the k-th TS, the signal reflected and amplified by the active IRS is given by

$$\mathbf{y}_{r,k} = \mathbf{\Phi}_k \mathbf{h}_{r,k} \sqrt{p_k} s_k + \mathbf{\Phi}_k \mathbf{n}_r, \quad k \in \mathcal{K}, \tag{1}$$

where $\mathbf{n}_r \in \mathbb{C}^{N \times 1}$ represents the thermal noise generated at the active IRS, which is distributed as $\mathcal{CN}\left(0, \sigma_r^2 \mathbf{I}_N\right)$. Note that the active IRS amplifies both the incident signal and noise. We denote the maximum amplification power of the active IRS as P_r and thus we have [31]

$$\mathrm{E}\left[\left\|\mathbf{y}_{r,k}\right\|^2\right] = p_k \left\|\mathbf{\Phi}_k \mathbf{h}_{r,k}\right\|^2 + \sigma_r^2 \left\|\mathbf{\Phi}_k\right\|_F^2 \leq P_r, \quad k \in \mathcal{K}. \tag{2}$$

Additionally, the received signal at the AP in the k-th TS can be expressed as

$$\mathbf{y}_k = \left(h_{d,k_l} + \mathbf{g}^H \mathbf{\Phi}_k \mathbf{h}_{r,k}\right) \sqrt{p_k} s_k + \mathbf{g}^H \mathbf{\Phi}_k \mathbf{n}_r + n, \quad k \in \mathcal{K}, \tag{3}$$

where $n \sim \mathcal{CN}\left(0, \sigma^2\right)$ denotes the additive white Gaussian noise at the AP. As such, the sum achievable throughput of the system in bits/Hz is given by

$$R_{\mathrm{TDMA}} = \sum_{k=1}^{K} \tau_k \log_2 \left(1 + \frac{p_k \left|h_{d,k} + \mathbf{g}^H \mathbf{\Phi}_k \mathbf{h}_{r,k}\right|^2}{\sigma^2 + \sigma_r^2 \left\|\mathbf{g}^H \mathbf{\Phi}_k\right\|^2}\right). \tag{4}$$

2) NOMA-Based Scheme: For the NOMA-based scheme, all the devices transmit simultaneously to the AP. As such, a common IRS beamforming pattern, denoted by $\mathbf{\Phi}$, is shared by all devices. In this case, the signal reflected and amplified by the active IRS is given by

$$\mathbf{y}_r = \mathbf{\Phi} \sum_{k=1}^{K} \mathbf{h}_{r,k} \sqrt{p_k} s_k + \mathbf{\Phi} \mathbf{n}_r. \tag{5}$$

Correspondingly, we have the following IRS amplification power constraint for the NOMA scheme

$$\mathrm{E}\left[\left\|\mathbf{y}_r\right\|^2\right] = \sum_{k=1}^{K} p_k \left\|\mathbf{\Phi} \mathbf{h}_{r,k}\right\|^2 + \sigma_r^2 \left\|\mathbf{\Phi}\right\|_F^2 \leq P_r. \tag{6}$$

The received signal at the AP can be further expressed as

$$\mathbf{y} = \sum_{k=1}^{K} \left(h_{d,k} + \mathbf{g}^H \mathbf{\Phi} \mathbf{h}_{r,k}\right) \sqrt{p_k} s_k + \mathbf{g}^H \mathbf{\Phi} \mathbf{n}_r + n. \tag{7}$$

To mitigate the multiuser interference, successive interference cancellation is performed at the AP. Taking device k as an example, the AP will first decode the message of device i, $\forall i < k$, before decoding the message of device k. Then, the message of device i, $\forall i < k$, will be subtracted from the composite signal. The message received from device i, $\forall i > k$, is treated as noise. Thus, the achievable sum throughput of all devices can be written as [35]

$$R_{\mathrm{NOMA}} = \tau \log_2 \left(1 + \sum\nolimits_{k=1}^{K} \frac{p_k \left|h_{d,k} + \mathbf{g}^H \mathbf{\Phi} \mathbf{h}_{r,k}\right|^2}{\sigma^2 + \sigma_r^2 \left\|\mathbf{g}^H \mathbf{\Phi}\right\|^2}\right), \tag{8}$$

where τ denotes the transmission time duration for all devices and $\tau \leq T_{\max}$.

C. Problem Formulation

We aim to maximize the sum throughput[2] of the considered system by jointly optimizing the time allocation, the transmit power of each device, and the active IRS beamforming. For the TDMA-based scheme, the optimization problem can be

[2]The device fairness issue for TDMA can be addressed by reformulating its associated optimization problem as a weighted sum throughput maximization one. By varying the values of these weights, the system designer is able to set different priorities and enforce certain notions of fairness among devices. Moreover, it can be checked that the weights do not affect the algorithm design in Section III-A. Therefore, we naturally assume that all the devices are equally weighted in this paper without loss of generality.

expressed as

$$\max_{\{\tau_k\},\{p_k\},\{\mathbf{\Phi}_k\}} \quad \sum_{k=1}^{K}\tau_k\log_2\left(1+\frac{p_k\left|h_{d,k}+\mathbf{g}^H\mathbf{\Phi}_k\mathbf{h}_{r,k}\right|^2}{\sigma^2+\sigma_r^2\left\|\mathbf{g}^H\mathbf{\Phi}_k\right\|^2}\right) \tag{9a}$$

$$\text{s.t.} \quad \tau_k p_k \le E_k, \quad \forall k\in\mathcal{K}, \tag{9b}$$

$$\sum_{k=1}^{K}\tau_k \le T_{\max}, \tag{9c}$$

$$\tau_k \ge 0,\ p_k \ge 0, \quad \forall k\in\mathcal{K}, \tag{9d}$$

$$p_k\left\|\mathbf{\Phi}_k\mathbf{h}_{r,k}\right\|^2+\sigma_r^2\left\|\mathbf{\Phi}_k\right\|_F^2\le P_r, \quad \forall k\in\mathcal{K}. \tag{9e}$$

For problem (9), constraint (9b) ensures that the total energy consumed at each device cannot exceed its available energy. Constraints (9c) and (9d) are the total transmission time constraint and the non-negative constraints on the optimization variables, respectively, while constraint (9e) indicates that the amplification power of the active IRS should not exceed the maximum allowed power. Similarly, the optimization problem associated with the NOMA-based scheme can be formulated as[3]

$$\max_{\tau,\{p_k\},\mathbf{\Phi}} \quad \tau\log_2\left(1+\sum_{k=1}^{K}\frac{p_k\left|h_{d,k}+\mathbf{g}^H\mathbf{\Phi}\mathbf{h}_{r,k}\right|^2}{\sigma^2+\sigma_r^2\left\|\mathbf{g}^H\mathbf{\Phi}\right\|^2}\right) \tag{10a}$$

$$\text{s.t.} \quad \tau p_k \le E_k, \quad \forall k\in\mathcal{K}, \tag{10b}$$

$$\tau \le T, \tag{10c}$$

$$\tau \ge 0,\ p_k \ge 0, \quad \forall k\in\mathcal{K}, \tag{10d}$$

$$\sum_{k=1}^{K} p_k\left\|\mathbf{\Phi}\mathbf{h}_{r,k}\right\|^2+\sigma_r^2\left\|\mathbf{\Phi}\right\|_F^2\le P_r. \tag{10e}$$

The above two problems (9) and (10) are all non-convex since the optimization variables are closely coupled in the objective function and constraints. Therefore, there are no standard methods for solving such non-convex optimization problems optimally in general. Moreover, it is worth noting that the hidden structures of problems (9) and (10) are fundamentally different. Specifically, in (10a) and (10e), the common IRS beamforming pattern, i.e., $\mathbf{\Phi}$, is coupled with all p_k's, while the transmit power of device k, i.e., p_k, is only coupled with its individual IRS beamforming pattern, i.e., $\mathbf{\Phi}_k$, in (9a) and (9e). These issues have non-trivial effects on the algorithms design and performance comparison for the optimal values of problems (9) and (10). In the next section, by deeply exploiting inherent properties of the corresponding optimization problems, we propose efficient algorithms to obtain high-quality solutions for them and provide a fundamental performance comparison for the active IRS aided TDMA and NOMA schemes.

[3]Note that the fairness of devices for the uplink NOMA scheme can be guaranteed by allowing time-sharing among different decoding orders [42]. How to ensure the individual data rate for different devices is another interesting problem for active IRS-aided energy-constrained NOMA communication systems. Towards this end, it is very crucial to additionally design the decoding order, IRS beamforming vector, and transmit power control at these devices, which, however, is left for our future work.

III. Investigations on Active IRS Aided TDMA and NOMA Schemes

In this section, we study the active IRS aided TDMA and NOMA schemes. By exploiting the specific structures of the associated problems, two dedicated algorithms are proposed for problems (9) and (10), respectively. Furhermore, we provide a theoretical performance comparison for the achievable sum throughput of the active IRS aided TDMA and NOMA schemes.

A. Proposed Algorithm for Active IRS Aided TDMA

For the TDMA-based scheme, the sum throughput maximization problem can be rewritten in a more tractable form as

$$\max_{\{\tau_k\},\{p_k\},\{\mathbf{v}_k\}} \quad \sum_{k=1}^{K}\tau_k\log_2\left(1+\frac{p_k\left|h_{d,k}+\mathbf{v}_k^H\mathbf{q}_k\right|^2}{\sigma^2+\sigma_r^2\mathbf{v}_k^H\mathbf{G}\mathbf{v}_k}\right) \tag{11a}$$

$$\text{s.t.} \quad (9b),(9c),(9d), \tag{11b}$$

$$p_k\mathbf{v}_k^H\mathbf{H}_{r,k}\mathbf{v}_k+\sigma_r^2\|\mathbf{v}_k\|^2\le P_r, \quad \forall k\in\mathcal{K}, \tag{11c}$$

where $\mathbf{v}_k^H=[\phi_{k,1},\ldots,\phi_{k,N}]$, $\mathbf{q}_k=\operatorname{diag}\left(\mathbf{g}^H\right)\mathbf{h}_{r,k}$, $\mathbf{G}=\operatorname{diag}\left(\left|[\mathbf{g}]_1\right|^2,\ldots\left|[\mathbf{g}]_N\right|^2\right)$, and $\mathbf{H}_{r,k}=\operatorname{diag}\left(\left|[\mathbf{h}_{r,k}]_1\right|^2,\ldots\left|[\mathbf{h}_{r,k}]_N\right|^2\right)$. Generally, problem (11) is a non-convex optimization problem and difficult to solve optimally due to the non-concave objective function as well as coupled variables in constraints (9b) and (11c). Different from the conventional passive IRS-aided communication systems, the new IRS amplification power constraint, i.e., constraint (11c), is involved. As a result, the value of the amplification coefficient at each element for the active IRS may be reduced as the transmit power of each device, i.e., p_k, increases, which may weaken the effectiveness of the active IRS and thus has negative effects on the sum throughput. Therefore, it is not clear whether the energy of each device would be used up for maximizing the sum throughput, i.e., (9b) is active or not at the optimal solution, which motivates the following lemma.

Lemma 1: At the optimal solution of problem (11), constraint (9b) is always met with equality, i.e., $\tau_k p_k = E_k$.

Proof: Please refer to Appendix A. ■

Lemma 1 reveals that for the active IRS-aided energy-constrained uplink communication system with TDMA, each device will deplete all of its energy at the optimal solution, i.e., constraint (9b) holds with equality. Thus, problem (11) is equivalently simplified to the following

$$\max_{\{\tau_k\},\{\mathbf{v}_k\}} \quad \sum_{k=1}^{K}\tau_k\log_2\left(1+\frac{E_k\left|h_{d,k}+\mathbf{v}_k^H\mathbf{q}_k\right|^2}{\tau_k\left(\sigma^2+\sigma_r^2\mathbf{v}_k^H\mathbf{G}\mathbf{v}_k\right)}\right) \tag{12a}$$

$$\text{s.t.} \quad \tau_k \ge 0,\ \forall k\in\mathcal{K}, \tag{12b}$$

$$\frac{E_k}{\tau_k}\mathbf{v}_k^H\mathbf{H}_{r,k}\mathbf{v}_k+\sigma_r^2\|\mathbf{v}_k\|^2\le P_r, \quad \forall k\in\mathcal{K}. \tag{12c}$$

$$(9c). \tag{12d}$$

Note that by exploiting Lemma 1, constraint (11c) is transformed into (12c), which is a convex constraint since E_k is a constant. The remaining challenge for solving problem (12) is the non-concave objective function. Nevertheless, we propose an efficient algorithm to solve it, where all the variables are optimized simultaneously. First, we introduce a set of slack variables denoted by $\{S_k\}$ and reformulate problem (12) as

$$\max_{\{\tau_k\},\{\mathbf{v}_k\},\{S_k\}} \quad \sum_{k=1}^{K} \tau_k \log_2\left(1+\frac{S_k}{\tau_k}\right) \tag{13a}$$

$$\text{s.t.} \qquad S_k \le \frac{E_k\left|h_{d,k}+\mathbf{v}_k^H\mathbf{q}_k\right|^2}{\sigma^2+\sigma_r^2\mathbf{v}_k^H\mathbf{G}\mathbf{v}_k},\ \forall k\in\mathcal{K}, \tag{13b}$$

$$(9c),(12b),(12c). \tag{13c}$$

For the optimal solution of problem (13), constraint (13b) is met with equality strictly since we can always increase the objective value by increasing S_k until constraint (13b) becomes active. Thus, problem (13) is equivalent to problem (12). However, constraint (13b) is still non-convex. To deal with constraint (13b), we rewrite it into a more tractable form as

$$\sigma^2+\sigma_r^2\mathbf{v}_k^H\mathbf{G}\mathbf{v}_k \le \frac{E_k\left|h_{d,k}+\mathbf{v}_k^H\mathbf{q}_k\right|^2}{S_k}, \forall k\in\mathcal{K}. \tag{14}$$

It is observed that $\sigma^2+\sigma_r^2\mathbf{v}_k^H\mathbf{G}\mathbf{v}_k$ is a convex quadratic function of $\mathbf{v}_k$ while the right-hand-side of (14) is jointly convex with respect to $\mathbf{v}_k$ and S_k. Recall that any convex function is globally lower-bounded by its first-order Taylor expansion at any feasible point, which motivates us to employ the SCA technique to deal with the non-convexity of (14). Specifically, for the given local point $\left\{\mathbf{v}_k^{(l)}, S_k^{(l)}\right\}$ in the l-th iteration, we have the following lower bound as

$$\begin{aligned}&\frac{E_k\left|h_{d,k}+\mathbf{v}_k^H\mathbf{q}_k\right|^2}{S_k}\\ &\ge \frac{2E_k}{S_k^{(l)}}\mathrm{Re}\left(\left(h_{d,k}+\mathbf{v}_k^H\mathbf{q}_k\right)^H\left(h_{d,k}+\left(\mathbf{v}_k^{(l)}\right)^H\mathbf{q}_k\right)\right)\\ &\quad-\frac{E_k\left|h_{d,k}+\left(\mathbf{v}_k^{(l)}\right)^H\mathbf{q}_k\right|^2}{\left(S_k^{(l)}\right)^2}S_k \triangleq f_k^{lb}\left(\mathbf{v}_k,S_k\right).\end{aligned} \tag{15}$$

It can be readily checked that $f_k^{lb}(\mathbf{v}_k,S_k)$ is linear and convex with respect to $\{\mathbf{v}_k,S_k\}$. As such, with the lower bound in (15), constraint (13b) is transformed to

$$\sigma^2+\sigma_r^2\mathbf{v}_k^H\mathbf{G}\mathbf{v}_k \le f_k^{lb}\left(\mathbf{v}_k,S_k\right), \forall k\in\mathcal{K}. \tag{16}$$

Then, problem (13) can be approximated by

$$\max_{\{\tau_k\},\{\mathbf{v}_k\},\{S_k\}} \quad \sum_{k=1}^{K} \tau_k \log_2\left(1+\frac{S_k}{\tau_k}\right) \tag{17a}$$

$$\text{s.t.} \qquad (9c),(12c),(12d),(16), \tag{17b}$$

which is a convex optimization problem. Thus, we can apply existing standard convex optimization tools to successively solve it optimally until the convergence is achieved. After convergence, it can be guaranteed that a locally optimal solution for original problem (11) can be obtained. The computational complexity of this algorithm lies in solving problem (17) and is given by $\mathcal{O}\left((K(N+2))^{3.5}I_{\mathrm{Iter}}\right)$, where I_{Iter} represents the number of iterations required for convergence.

B. Proposed Algorithm for NOMA

For the active IRS-aided NOMA scheme, the sum throughput maximization problem can be rewritten as

$$\max_{\tau,\{p_k\},\mathbf{v}} \quad \tau\log_2\left(1+\frac{\sum_{k=1}^{K}p_k\left|h_{d,k}+\mathbf{v}^H\mathbf{q}_k\right|^2}{\sigma^2+\sigma_r^2\mathbf{v}^H\mathbf{G}\mathbf{v}}\right) \tag{18a}$$

$$\text{s.t.} \qquad (10b),(10c),(10d) \tag{18b}$$

$$\sum_{k=1}^{K} p_k\mathbf{v}^H\mathbf{H}_{r,k}\mathbf{v}+\sigma_r^2\|\mathbf{v}\|^2 \le P_r, \tag{18c}$$

where $\mathbf{v}^H=[\phi_1,\ldots,\phi_N]$. Different from problem (11), all devices share the same IRS beamforming vector in problem (18) and the common IRS beamforming vector $\mathbf{v}$ is coupled with the transmit power of all devices in (18a) and (18c). Thus, $\mathbf{v}$ cannot be flexibly adjusted for each individual device and the amplification gains of the active IRS may be locked if all devices transmit at their maximum allowed power simultaneously. Furthermore, different from the TDMA case, some devices may not use up all of their energy at the optimal solution for the NOMA case. As such, the algorithm proposed for problem (11) is not applicable to the more challenging problem (18), which thus calls for the new algorithm design. Problem (18) is generally intractable due to the non-concave objective function and non-convex constraints (10b) and (18c). To tackle the coupled variables p_k and τ in constraint (10b), we have the following lemma.

Lemma 2: At the optimal solution of problem (18), constraint (10c) is strictly met with equality, i.e., $\tau=T_{\max}$.

Proof: We firstly define $e_k=p_k\tau$ and then the objective function of problem (18) can be rewritten as

$$\tau\log_2\left(1+\frac{\sum_{k=1}^{K}e_k\left|h_{d,k}+\mathbf{v}^H\mathbf{q}_k\right|^2}{\tau\left(\sigma^2+\sigma_r^2\mathbf{v}^H\mathbf{G}\mathbf{v}\right)}\right). \tag{19}$$

Suppose τ^* is the optimal transmission time. Then, we show that $\tau^*=T_{\max}$ by contradiction as follows. Assume that $\Xi^*=\{\tau^*,p_k^*,\mathbf{v}^*\}$ achieves the optimal solution of problem (18) and $\tau^*<T_{\max}$. Then, we construct a different solution $\tilde{\Xi}=\{\tilde{\tau},\tilde{p}_k,\tilde{\mathbf{v}}\}$, where $\tau^*<\tilde{\tau}\le T_{\max}$, $e_k^*=\tilde{\tau}\tilde{p}_k=\tau^*p_k^*$, and $\tilde{\mathbf{v}}=\mathbf{v}^*$. It can be readily verified that $\{\tilde{p}_k,\tilde{\mathbf{v}}\}$ satisfies constraint (18c) since $\tilde{p}_k\le p_k^*$. As such, the constructed solution satisfies all the constraints in problem (18). Since (19) is an increasing function with respect to τ and $\tilde{\tau}>\tau^*$, we have

$$\begin{aligned}&\tilde{\tau}\log_2\left(1+\frac{\sum_{k=1}^{K}e_k^*\left|h_{d,k}+\tilde{\mathbf{v}}^H\mathbf{q}_k\right|^2}{\tilde{\tau}\left(\sigma^2+\sigma_r^2\tilde{\mathbf{v}}^H\mathbf{G}\tilde{\mathbf{v}}\right)}\right)\\ &\qquad>\tau^*\log_2\left(1+\frac{\sum_{k=1}^{K}e_k^*\left|h_{d,k}+\mathbf{v}^{*H}\mathbf{q}_k\right|^2}{\tau^*\left(\sigma^2+\sigma_r^2\mathbf{v}^{*H}\mathbf{G}\mathbf{v}^*\right)}\right).\end{aligned} \tag{20}$$

(20) indicates that the constructed solution $\tilde{\Xi}$ achieves a higher objective value, which contradicts that the optimal $\tau^* < T_{\max}$. This thus completes the proof. ■

Exploiting Lemma 2, problem (18) can be simplified to

$$\max_{\{p_k\},\mathbf{v}} \quad T_{\max}\log_2\left(1+\frac{\sum_{k=1}^{K} p_k\left|h_{d,k}+\mathbf{v}^H\mathbf{q}_k\right|^2}{\sigma^2+\sigma_r^2\mathbf{v}^H\mathbf{G}\mathbf{v}}\right) \tag{21a}$$

$$\text{s.t.} \quad p_k \le \frac{E_k}{T_{\max}}, \quad \forall k\in\mathcal{K}, \tag{21b}$$

$$p_k \ge 0, \quad \forall k\in\mathcal{K}, \tag{21c}$$

$$(18c). \tag{21d}$$

The main challenges for solving problem (21) are the non-concave objective function (21a) and the coupled variables $\{p_k, \mathbf{v}\}$ involved in constraint (21d), which motivates us to apply the AO-based method to solve it. Specifically, we divide all the variables into two blocks, i.e., 1) IRS beamforming vector $\mathbf{v}$, and 2) power control p_k, and then each block of variables is optimized in an iterative way, until convergence is achieved.

1) IRS Beamforming Optimization: For any given transmit power p_k, the IRS beamforming vector optimization problem is given by

$$\max_{\mathbf{v}} \quad T_{\max}\log_2\left(1+\frac{\sum_{k=1}^{K} p_k\left|h_{d,k}+\mathbf{v}^H\mathbf{q}_k\right|^2}{\sigma^2+\sigma_r^2\mathbf{v}^H\mathbf{G}\mathbf{v}}\right) \tag{22a}$$

$$\text{s.t.} \quad (21d). \tag{22b}$$

Let $\left|h_{d,k}+\mathbf{v}^H\mathbf{q}_k\right|^2 = \left|\bar{\mathbf{v}}^H\bar{\mathbf{q}}_k\right|^2$, where $\bar{\mathbf{v}}^H = \left[\mathbf{v}^H, 1\right]$ and $\bar{\mathbf{q}}_k = \left[\mathbf{q}_k^H, h_{d,k}^H\right]^H$. Define $\bar{\mathbf{Q}}_k = \bar{\mathbf{q}}_k\bar{\mathbf{q}}_k^H$, $\bar{\mathbf{V}} = \bar{\mathbf{v}}\bar{\mathbf{v}}^H$, which needs to satisfy $\bar{\mathbf{V}}\succeq\mathbf{0}$, $\mathrm{rank}\left(\bar{\mathbf{V}}\right)=1$ and $\left[\bar{\mathbf{V}}\right]_{N+1,N+1}=1$. We thus have

$$\sum_{k=1}^{K} p_k\left|h_{d,k}+\mathbf{v}^H\mathbf{q}_k\right|^2 = \sum_{k=1}^{K} p_k\mathrm{Tr}\left(\bar{\mathbf{Q}}_k\bar{\mathbf{V}}\right) = \mathrm{Tr}\left(\bar{\mathbf{Q}}\bar{\mathbf{V}}\right), \tag{23}$$

where $\bar{\mathbf{Q}} = \sum_{k=1}^{K} p_k\bar{\mathbf{Q}}_k$. Define $\bar{\mathbf{G}} = \mathrm{diag}\left(\mathbf{G}, 0\right)$, $\bar{\mathbf{H}}_{r,k} = \mathrm{diag}\left(\mathbf{H}_{r,k}, 0\right)$, and $\bar{\mathbf{H}}_r = \sum_{k=1}^{K} p_k\bar{\mathbf{H}}_{r,k} + \sigma_r^2\mathbf{\Pi}_r$, where $\mathbf{\Pi}_r = \mathrm{diag}\left(1,\ldots 1, 0\right)$. We further have

$$\sigma^2+\sigma_r^2\mathbf{v}^H\mathbf{G}\mathbf{v} = \sigma^2+\sigma_r^2\mathrm{Tr}\left(\bar{\mathbf{G}}\bar{\mathbf{V}}\right), \tag{24}$$

$$p_k\mathbf{v}^H\mathbf{H}_{r,k}\mathbf{v}+\sigma_r^2\|\mathbf{v}\|^2 = \mathrm{Tr}\left(\bar{\mathbf{H}}_r\bar{\mathbf{V}}\right). \tag{25}$$

As such, we can reformulate problem (22) in an equivalent form as follows

$$\max_{\bar{\mathbf{V}}} \quad \frac{\mathrm{Tr}\left(\bar{\mathbf{Q}}\bar{\mathbf{V}}\right)}{\sigma^2+\sigma_r^2\mathrm{Tr}\left(\bar{\mathbf{G}}\bar{\mathbf{V}}\right)} \tag{26a}$$

$$\text{s.t.} \quad \mathrm{Tr}\left(\bar{\mathbf{H}}_r\bar{\mathbf{V}}\right) \le P_r, \tag{26b}$$

$$\bar{\mathbf{V}}\succeq\mathbf{0}, \tag{26c}$$

$$\left[\bar{\mathbf{V}}\right]_{N+1,N+1}=1, \tag{26d}$$

$$\mathrm{rank}\left(\bar{\mathbf{V}}\right)=1. \tag{26e}$$

Problem (26) is still intractable due to the fraction form in the objective function and the rank-one constraint. To tackle this issue, we relax the rank-one constraint and apply the Charnes-Cooper transformation to reformulate it as a linear form as

$$\max_{\bar{\mathbf{V}},t} \quad \mathrm{Tr}\left(\bar{\mathbf{Q}}\bar{\mathbf{V}}\right) \tag{27a}$$

$$\text{s.t.} \quad \sigma_r^2\mathrm{Tr}\left(\bar{\mathbf{G}}\bar{\mathbf{V}}\right)+t\sigma^2=1, \tag{27b}$$

$$\mathrm{Tr}\left(\bar{\mathbf{H}}_r\bar{\mathbf{V}}\right) \le tP_r, \tag{27c}$$

$$t>0, \tag{27d}$$

$$\left[\bar{\mathbf{V}}\right]_{N+1,N+1}=t, \quad (26c). \tag{27e}$$

Lemma 3: By relaxing the rank-one constraint (37e) in problem (26), problem (26) is equivalent to problem (27).

Proof: First, given any feasible solution $\left\{\bar{\mathbf{V}}\right\}$ to problem (26), it can be verified that with the solution $\left\{\bar{\mathbf{V}}/\left(\sigma^2+\sigma_r^2\mathrm{Tr}\left(\bar{\mathbf{G}}\bar{\mathbf{V}}\right)\right), 1/\left(\sigma^2+\sigma_r^2\mathrm{Tr}\left(\bar{\mathbf{G}}\bar{\mathbf{V}}\right)\right)\right\}$, problem (27) achieves the same objective value as that of (26). Then, given any feasible solution $\left\{\bar{\mathbf{V}}, t\right\}$ to problem (27), it can be similarly demonstrated that with the solution $\left\{\bar{\mathbf{V}}/t\right\}$, problem (26) achieves the same objective value. As such, problem (26) and (27) have the same optimal value. Lemma 3 is thus proved. ■

It can be readily verified that problem (27) is a convex optimization problem, whose optimal solution can be efficiently solved by the standard convex optimization techniques. According to Lemma 3, we can solve problem (27) instead of solving (26).

Remark 1: According to Theorem 3.2 in [43], we can conclude that there always exists a optimal solution $\bar{\mathbf{V}}^*$ to problem (27), satisfies the following constraint:

$$\mathrm{Rank}^2\left(\bar{\mathbf{V}}^*\right)+\mathrm{Rank}^2\left(t^*\right)\le 3. \tag{28}$$

We note that $\mathrm{Rank}^2\left(t^*\right)=1$. Thus, we have $\mathrm{Rank}\left(\bar{\mathbf{V}}^*\right)=1$. There always exists rank-one solution for problem (27).

Based on Remark 1, we can obtain the optimal rank-one solution for problem (27), denoted by $\left\{\bar{\mathbf{V}}^*, t^*\right\}$. Thus, the optimal solution for problem (26) is obtained as $\bar{\mathbf{V}}^*/t^*$. By performing singular value decomposition (SVD) for $\bar{\mathbf{V}}^*/t^*$, the optimal solution $\mathbf{v}^*$ can be found for original problem (22).

2) Power Control Optimization: For any given IRS beamforming vector $\mathbf{v}$, the power control optimization problem is given by

$$\max_{\{p_k\}} \quad \frac{\sum_{k=1}^{K} p_k\left|h_{d,k}+\mathbf{v}^H\mathbf{q}_k\right|^2}{\sigma^2+\sigma_r^2\mathbf{v}^H\mathbf{G}\mathbf{v}} \tag{29a}$$

$$\text{s.t.} \quad (21b), (21c), (21d). \tag{29b}$$

Since the objective function and all constraints in (29) are linear, problem (29) is a convex optimization problem, which can be optimally solved by the standard convex optimization methods.

3) Overall Algorithm and Computational Complexity Analysis: Based on the solutions to the above two subproblems, an efficient AO algorithm is proposed, where the IRS beamforming vector and power control are alternately optimized until convergence is achieved. Note that the objective value of problem (21) is non-decreasing by alternately optimizing $\{\mathbf{v}\}$ and $\{p_k\}$, thus the proposed AO algorithm is guaranteed

to converge. The mainly computational complexity of this AO algorithm lies from solving problems (27) and (29). Specifically, the computational complexity for solving problems (27) and (29) is given by $\mathcal{O}\left((N+1)^{3.5}\right)$ and $\mathcal{O}\left(K^{3.5}\right)$, respectively. Therefore, the total complexity of the AO algorithm is $\mathcal{O}\left(\left((N+1)^{3.5}+K^{3.5}\right)I_{\mathrm{AO}}\right)$, where I_{AO} denotes the number of iterations required to reach convergence.

Remark 2: Note that the AO design principles are also applicable to the problem formulation with the individual signal-to-interference-plus-noise ratio (SINR) constraint of each device. Without loss of generality, for device k, it is assumed that the AP will first decode the message of device i, $\forall i < k$, before decoding the message of device k. Then, the message of device i, $\forall i < k$, will be subtracted from the received composite signal. Meanwhile, the offloading message received from device i, $\forall i > k$, is treated as noise. As such, under the SDP transformation, each device's individual SINR constraint can be expressed as

$$\frac{p_k \mathrm{Tr}\left(\bar{\mathbf{Q}}_k \bar{\mathbf{V}}\right)}{\sum_{i=k+1}^{K} p_i \mathrm{Tr}\left(\bar{\mathbf{Q}}_i \bar{\mathbf{V}}\right) + \sigma^2 + \sigma_r^2 \mathrm{Tr}\left(\bar{\mathbf{G}} \bar{\mathbf{V}}\right)} \geq \gamma_k, \quad \forall k, \quad (30)$$

where γ_k is minimum SINR required by device k. Constraint (30) can be equivalently transformed into

$$p_k \mathrm{Tr}\left(\bar{\mathbf{Q}}_k \bar{\mathbf{V}}\right) \geq \gamma_k \left(\sum_{i=k+1}^{K} p_i \mathrm{Tr}\left(\bar{\mathbf{Q}}_i \bar{\mathbf{V}}\right) + \sigma^2 + \sigma_r^2 \mathrm{Tr}\left(\bar{\mathbf{G}} \bar{\mathbf{V}}\right) \right), \quad \forall k. \quad (31)$$

It can be easily verified that constraint (31) is convex with respect to $\bar{\mathbf{V}}$ ($\{p_k\}$) when $\{p_k\}$ ($\bar{\mathbf{V}}$) is fixed. Thus, the principle of the proposed AO algorithm in this paper can also be applicable to solve such a problem with each device's individual SINR constraint.

C. Active IRS Aided TDMA Versus NOMA

Compared to TDMA, NOMA is expected to achieve a higher throughput in a conventional communication system by allowing multiple devices simultaneously to access the same spectrum. In our considered active IRS aided energy-constrained IoT systems, each device can occupy its dedicated IRS beamforming under the TDMA-based scheme, while all the devices share the same IRS beamforming vector under the NOMA-based scheme. As such, it is not clear which MA scheme can achieve higher throughput. In this subsection, we give some discussions about the performance comparison for the active IRS aided TDMA and NOMA schemes.

To provide a comprehensive comparison for the active IRS-aided TDMA and NOMA schemes, we introduce a special case of the active IRS aided TDMA where only one beamforming vector is available for assisting uplink transmission, which leads to the following problem formulation:

$$\max_{\{\tau_k\},\{p_k\},\mathbf{v}} \quad \sum_{k=1}^{K} \tau_k \log_2 \left(1 + \frac{p_k \left|h_{d,k} + \mathbf{v}^H \mathbf{q}_k\right|^2}{\sigma^2 + \sigma_r^2 \mathbf{v}^H \mathbf{G} \mathbf{v}}\right) \quad (32a)$$

$$\text{s.t.} \quad (9b), (9c), (9d), \quad (32b)$$

$$p_k \mathbf{v}^H \mathbf{H}_{r,k} \mathbf{v} + \sigma_r^2 \|\mathbf{v}\|^2 \leq P_r, \quad \forall k \in \mathcal{K}. \quad (32c)$$

By replacing $\mathbf{v}_k$ by $\mathbf{v}$ in (15), our proposed SCA-based algorithm in Section III-A can be directly extended to obtain the suboptimal solution of problem (32). Intuitively, it seems that constraint (18c) is tighter than (32c) in terms of the transmit power. Thus, it may be expected that the optimal value of problem (32) is no smaller than that of (18). However, the result is counterintuitive as shown in the following proposition.

Proposition 1: Denote the optimal values of problem (18) and (32) by R^*_{NOMA} and $R^{(lb)*}_{\mathrm{TDMA}}$, respectively, it follows that

$$R^*_{\mathrm{NOMA}} \geq R^{(lb)*}_{\mathrm{TDMA}}. \quad (33)$$

If

$$\tilde{p}_k \left|h_{d,k} + \tilde{\mathbf{v}}^H \mathbf{q}_k\right|^2 \neq \tilde{p}_j \left|h_{d,j} + \tilde{\mathbf{v}}^H \mathbf{q}_j\right|^2, \exists k, j \in \mathcal{K}, \quad (34)$$

$R^*_{\mathrm{NOMA}} > R^{(\mathrm{lb})*}_{\mathrm{TDMA}}$ holds, where $\{\tilde{p}_k, \tilde{\mathbf{v}}\}$ is the optimal solution of problem (32).

Moreover, a sufficient condition for $R^*_{\mathrm{NOMA}} = R^{(lb)*}_{\mathrm{TDMA}}$ is

$$\breve{p}_k \breve{\mathbf{v}}^H \mathbf{H}_{r,k} \breve{\mathbf{v}} + \sigma_r^2 \left\|\breve{\mathbf{v}}\right\|^2 \leq P_r, \forall k \in \mathcal{K}, \quad (35)$$

where $\left\{\breve{p}_k, \breve{\mathbf{v}}\right\}$ is the optimal solution of problem (32) when relaxing constraint (32c).

Proof: Please refer to Appendix B. ∎

Proposition 1 implies that the active IRS aided NOMA scheme achieves a no smaller sum throughput than its counterpart of the TDMA scheme when only one beamforming vector is available. This seems contradictory to the conclusion of the previous work [30] in the passive IRS aided energy-constrained scenario. The reason is that for an energy-constrained system, the transmit power of each device under the NOMA case is generally lower than that of the TDMA case. This makes NOMA be able to reap larger amplification gains of the active IRS compared to TDMA, which compensates the performance loss induced by the low transmit power. Based on Proposition 1, we give a sufficient condition for that the active IRS aided TDMA scheme can achieve a higher throughput than NOMA in the following theorem.

Theorem 1: Denote the optimal value of problem (11) by R^*_{TDMA}. Then, (35) serves as a sufficient condition for $R^*_{\mathrm{TDMA}} \geq R^*_{\mathrm{NOMA}}$.

Proof: According to the results in Proposition 1, we have $R^{(\mathrm{lb})*}_{\mathrm{TDMA}} = R^*_{\mathrm{NOMA}}$ when (35) is satisfied. Note that problem (32) is a special case of problem (11) with $\mathbf{v}_k = \mathbf{v}_j, \forall k, j \in \mathcal{K}$. As such, the optimal solution of problem (32) is also one feasible solution of problem (11), which yields $R^*_{\mathrm{TDMA}} \geq R^{(\mathrm{lb})*}_{\mathrm{TDMA}} = R^*_{\mathrm{NOMA}}$. This thus completes the proof. ∎

Note that considering the practical issue of the imperfect SIC would definitely cause the performance loss for the NOMA scheme. Our analytical result in Theorem 1 is still applicable to the practical scenario with imperfect SIC.

Remark 3: Theorem 1 implies that the active IRS aided TDMA scheme generally outperforms its counterpart with NOMA due to more IRS beamforming vectors can be

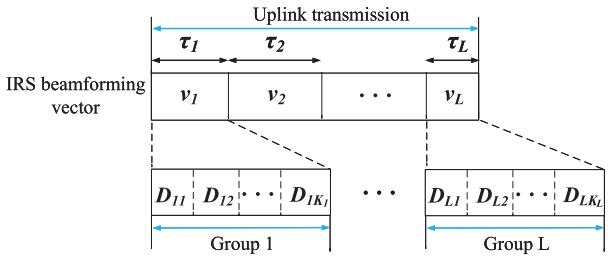


Fig. 2. The proposed general TDMA-NOMA scheme for active IRS-aided uplink communications.

exploited for TDMA. The above two MA schemes strike a balance between the throughput performance and the number of optimization variables as well as the feedback signalling overhead. Specifically, the TDMA-based scheme requires the AP to optimize and feedback KN IRS reflection coefficients (including both amplitudes and phase shifts) to the IRS, which increases linearly with the number of devices, while these required for the NOMA-based scheme is N. Therefore, the signaling overhead for TDMA in such an overloaded scenario is much heavier than that of NOMA. This motivates us to propose a more flexible scheme to fully harness the maximum performance gain of the active IRS with controllable signaling overhead and complexity, elaborated in the next section.

IV. An Overhead-Aware Hybrid TDMA-NOMA Scheme

Motivated by the discussions in the previous section, we develop a more general hybrid TDMA-NOMA-based uplink transmission scheme in this section for fully exploiting the tradeoff between the throughput performance and signalling overhead.[4]

A. Problem Formulation for General Hybrid TDMA-NOMA Scheme

The detailed transmission protocol is illustrated in Fig. 2. Without loss of generality, we assume that the IRS beamforming vectors can be reconfigured $L-1$ times during the transmission period, corresponding to L available IRS beamforming vectors, i.e., $\mathbf{v}_l, l \in \mathcal{L} = \{1, \ldots L\}$. We further partition K devices into L disjoint groups equally and the set of devices in l-th group is denoted by $\mathcal{K}_l$ with size $K_l = K/L$. Furthermore, the devices in different groups transmit in different TSs, while the devices in the same group transmit simultaneously employing NOMA. Specifically, each device k_l, $k_l \in \mathcal{K}_l$, will transmit in the l-th TS with the aid of IRS beamforming vector $\mathbf{v}_l$ and the time duration for the l-th TS is denoted by τ_l. Accordingly, the system sum throughput maximization problem by jointly optimizing the IRS beamforming vector, the transmit power, the time allocation, and the device grouping, can be formulated as

$$\max_{\{\tau_l\},\{p_{k_l}\},\{\mathbf{v}_l\},\mathcal{K}_l} \quad \sum_{l=1}^{L} \tau_l \log_2\left(1+\sum_{k_l\in\mathcal{K}_l} \frac{p_{k_l}\left|h_{d,k_l}+\mathbf{v}_l^H\mathbf{q}_{k_l}\right|^2}{\sigma^2+\mathbf{v}_l^H\mathbf{G}\mathbf{v}_l}\right) \tag{36a}$$

$$\text{s.t.} \quad \tau_l p_{k_l} \le E_{k_l},\ \forall l \in \mathcal{L}, \forall k_l \in \mathcal{K}_l, \tag{36b}$$

$$\sum_{l=1}^{L} \tau_l \le T_{\max}, \tag{36c}$$

$$\tau_l \ge 0,\ p_{k_l} \ge 0,\ \forall l \in \mathcal{L},\ \forall k_l \in \mathcal{K}_l, \tag{36d}$$

$$\sum\nolimits_{k_l\in\mathcal{K}_l} p_{k_l}\mathbf{v}_l^H\mathbf{H}_{r,k_l}\mathbf{v}_l + \sigma_r^2\|\mathbf{v}_l\|^2 \le P_r,\ \forall k_l \in \mathcal{K}_l. \tag{36e}$$

It is worth noting that problem (36) is equivalent to (9) when $L = K$. In addition, when $L = 1$, it can be verified that problem (36) is equivalent to (10). As such, both standalone TDMA and NOMA-based schemes are special cases of the proposed hybrid TDMA-NOMA-based scheme. By controlling the number groups, i.e., L, the proposed hybrid MA scheme can be applicable for any given number of IRS beamforming vectors imposed by the practical systems and operate in an overhead-aware manner.

It has been shown in [44] that the user grouping plays a critical role in determining the overall uplink hybrid NOMA system performance. In a hybrid NOMA system without IRS, the user grouping can be performed based on their channel conditions. However, for our considered active IRS-aided energy-constraint uplink systems, the device grouping becomes more complicated, as the concatenated channel quality, namely $p_{k_l}\left|h_{d,k_l}+\mathbf{v}_l^H\mathbf{q}_{k_l}\right|^2 \Big/ \left(\sigma^2+\mathbf{v}_l^H\mathbf{G}\mathbf{v}_l\right)$ depends not only on h_{d,k_l} but also on the coupled optimization variables (i.e., $\mathbf{v}_l$ and p_{k_l}). The highly challenging device grouping design controlled by the IRS beamforming vectors makes problem (36) intractable. To make it tractable, we handle this problem in two steps, i.e., obtaining the device grouping set and solving a throughput maximization problem under the obtained device grouping set. It is worth noting that the optimal device grouping can be obtained through the exhaustive searching. Unfortunately, this method is exponential with respect to the number of devices and is computationally prohibitive in an overloaded scenario.

B. Device Grouping Strategy

To further reduce the computational complexity of the exhaustive searching, we develop a computationally efficient device grouping method by taking into account both the combined channel quality and the available energy of each device. The key idea is to decouple the devices' influence on each other in the same group. Specifically, we introduce the concept of the trading signal-to-noise ratio (SNR) to measure the equivalent channel conditions of all devices. The trading SNR of device k is mathematically defined as

$$\eta_k = \frac{E_k\left|h_{d,k}+\mathbf{v}_k^H\mathbf{q}_k\right|^2}{T_{\max}\left(\sigma^2+\mathbf{v}_k^H\mathbf{G}\mathbf{v}_k\right)}, \forall k \in \mathcal{K}. \tag{37}$$

[4]The signalling overhead in this paper specifically refers to the feedback overhead incurred by sending the IRS reflection coefficients from the AP to the IRS.

Then, the optimal values of all devices' trading SNR can be obtained by solving K optimization problems in parallel as follows

$$\max_{\mathbf{v}_k} \quad \frac{E_k\left|h_{d,k}+\mathbf{v}_k^H\mathbf{q}_k\right|^2}{T_{\max}\left(\sigma^2+\mathbf{v}_k^H\mathbf{G}\mathbf{v}_k\right)} \tag{38a}$$

$$\text{s.t.} \quad \frac{E_k}{T}\mathbf{v}_k^H\mathbf{H}_{r,k}\mathbf{v}_k+\sigma_r^2\|\mathbf{v}_k\|^2\le P_r. \tag{38b}$$

Note that problem (38) can be solved optimally by applying Charnes-Cooper transformation-based SDP technique, which is omitted here for brevity. The optimal value for device k's trading SNR is denoted by η_k^*. Then, we order K devices in the ascending order of the optimal target SNR η_k^*. Without loss of generality, we assume that $\eta_1^*\le\eta_2^*\le\ldots\le\eta_K^*$. The results in [44] indicate that maintaining the smallest overall channel strength difference among different device groups is beneficial for improving the sum throughput of a hybrid NOMA system. As such, the device grouping is performed based on the criterion of letting the trading SNR differences among the devices in the same group as large as possible. We obtain the device groups as follows

$$\mathcal{K}_l=\left\{u_l,u_{l+L},\ldots u_{l+(K_l-1)L}\right\},l\in\mathcal{L} \tag{39}$$

where u_k represents device k.

C. Optimization Under Given Device Grouping

Note that the non-convex problem (36) is more challenging to solve than problems (11) and (18). Specifically, different from problem (11) that each device can occupy a dedicated IRS beamforming vector or problem (18) that all devices share the same IRS beamforming vector. As such, Lemma 1 and Lemma 2 can not be applied here to simplify the problem. To handle this issue, we extend the AO-based framework to solve such a problem by partitioning the entire optimization variables into two blocks, but with the different elements in each block, namely, $\{\tau_l,p_{k_l}\}$ and $\{\mathbf{v}_l\}$, for updating alternatingly. Next, we respectively solve the above two blocks.

1) Optimizing $\{\tau_l,p_{k_l}\}$: For the given block $\{\mathbf{v}_l\}$, the subproblem for optimizing $\{\tau_l,p_{k_l}\}$ is still a non-convex optimization problem due to the coupled variables τ_l and p_{k_l} in constraint (36b) and the non-concave objective function. We apply a change of variables as $e_{k_l}=\tau_l p_{k_l}$ and rewrite the subproblem in an equivalent form as follows

$$\max_{\{\tau_l\},\{e_{k_l}\}} \quad \sum_{l=1}^{L}\tau_l\log_2\left(1+\sum\nolimits_{k_l\in K_l}\frac{e_{k_l}\left|h_{d,k_l}+\mathbf{v}_l^H\mathbf{q}_{k_l}\right|^2}{\tau_l\left(\sigma^2+\mathbf{v}_l^H\mathbf{G}\mathbf{v}_l\right)}\right) \tag{40a}$$

$$\text{s.t.} \quad e_{k_l}\le E_{k_l},\quad \forall l\in\mathcal{L},\ \forall k_l\in\mathcal{K}_l, \tag{40b}$$

$$\tau_l\ge 0,\ e_{k_l}\ge 0,\ \forall l\in\mathcal{L},\ \forall k_l\in\mathcal{K}_l, \tag{40c}$$

$$\sum\nolimits_{k_l\in\mathcal{K}_l}e_{k_l}\mathbf{v}_l^H\mathbf{H}_{r,k_l}\mathbf{v}_l+\sigma_r^2\tau_l\|\mathbf{v}_l\|^2\le P_r\tau_l,\quad \forall l\in\mathcal{L}, \tag{40d}$$

$$\text{(36c)}. \tag{40e}$$

It can be readily checked that objective function (40a) is concave and all constraints in problem (40) are convex. Therefore, problem (40) is convex and its optimal solution can be efficiently obtained by using standard convex optimization solvers such as CVX.

2) Optimizing $\{\mathbf{v}_l\}$: For the given $\{\tau_l,p_{k_l}\}$, it is observed that optimization variables with respect to different groups are separable in both the objective function and constraints. Thus, the resultant problem with respect to $\{\mathbf{v}_l\}$ can be addressed by solving L independent subproblems in parallel, each with only one single constraint. Specifically, for group l, the corresponding subproblem with respect to $\mathbf{v}_l$'s, $\forall l\in\mathcal{L}$, is reduced to

$$\max_{\mathbf{v}_l} \quad \sum_{k_l\in\mathcal{K}_l}\frac{p_{k_l}\left|h_{d,k_l}+\mathbf{v}_l^H\mathbf{q}_{k_l}\right|^2}{\sigma^2+\mathbf{v}_l^H\mathbf{G}\mathbf{v}_l} \tag{41a}$$

$$\text{s.t.} \quad \sum\nolimits_{k_l\in\mathcal{K}_l}p_{k_l}\mathbf{v}_l^H\mathbf{H}_{r,k_l}\mathbf{v}_l+\sigma_r^2\|\mathbf{v}_l\|^2\le P_r. \tag{41b}$$

Since problem (41) has the same form as that of problem (22), it can be similarly solved optimally by applying the Charnes-Cooper transformation-based SDP technique and the details are thus omitted for brevity.

Similar to the discussions in Section III-B, the computational complexity of the AO algorithm for solving problem (36) is given by $\mathcal{O}\left(\left(L^{3.5}(N+1)^{3.5}+(2L)^{3.5}\right)I_{\text{AO}}\right)$.

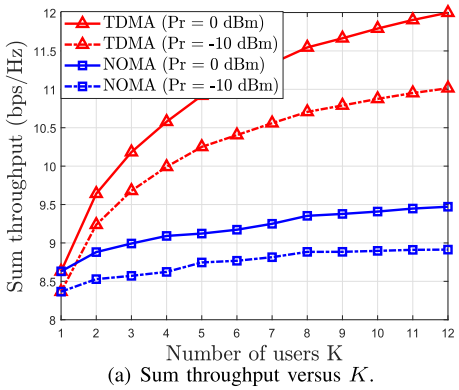
V. Numerical Results

In this section, we provide numerical results to validate the effectiveness of the proposed schemes and to draw useful insights into active IRS-aided energy-constrained IoT systems. The AP and the active IRS are located at $(0,0,0)$ meter (m) and $(x_{\text{IRS}},0,4)$ m, respectively, and the devices are uniformly and randomly distributed within a radius of 5 m centered at $(x_{\text{D}},0,4)$ m. The path-loss exponents of both the IRS-device and AP-IRS links are set to 2.2, while those of the AP-device links are set to 3.4 [26]. In addition, we assume that the AP-device link, the AP-IRS link, and the IRS-device link follow Rayleigh fading. The signal attenuation at a reference distance of 1 m is set as 30 dB. The other parameters are set as follows: $N=50$, $T=0.1$ s, $E_k=E_m,\forall m\ne k$, $x_{\text{IRS}}=0$, $x_{\text{D}}=30$ m, and $\sigma^2=\sigma_r^2=-75$ dBm [31].

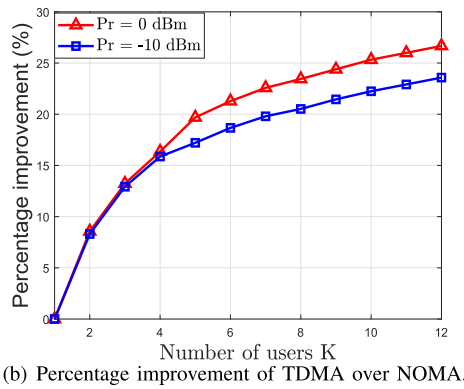
The initializations of our proposed algorithms are as follows. For the TDMA scheme, the initial time allocation, $\{\tau_k\}$, is obtained by letting $\tau_k=T_{\max}/K,\forall k\in\mathcal{K}$. Each element of the IRS beamforming vector for device k, $k\in\mathcal{K}$, is initialed as $[\mathbf{v}_k]_n=e^{j\left(\arg\left([\mathbf{q}_k]_n\right)+\arg(h_{d,k})\right)}$. For the NOMA scheme, the initial transmit power for each device, $\{p_k\}$, is obtained by letting $p_k=E_k/T_{\max},\forall k\in\mathcal{K}$ and then the initial IRS beamforming vector can be obtained by solving problem (22) under the initialized transmit power. Similarly, for the hybrid TDMA-NOMA scheme, the initial time allocation, $\{\tau_l\}$, and the initial transmit power, $\{p_k\}$, are obtained by letting $\tau_l=T_{\max}/L,\forall l\in\mathcal{L}$, and $p_k=E_kL/T_{\max},\forall k\in\mathcal{K}$. Then, the initial IRS beamforming vectors, $\{\mathbf{v}_l\}$, can be obtained by solving problem (41).

A. Performance Comparison for TDMA and NOMA-Based Schemes

In this subsection, we compare two types of MA schemes, i.e., TDMA and NOMA, in terms of the sum throughput



(a) Sum throughput versus K.



(b) Percentage improvement of TDMA over NOMA.

Fig. 3. Performance comparison for active IRS aided TDMA and NOMA.

in Fig. 3. Specifically, the sum throughput of the TDMA and NOMA schemes versus the number of devices K is presented in Fig. 3(a) under different values of P_r and the corresponding percentage improvement provided by TDMA over NOMA is presented in Fig. 3(b). For the TDMA-based scheme, we observe that the sum throughput increases as K increases. This is expected since scheduling one more device would always result in a higher sum throughput, which verifies our analysis in Lemma 1. While for the NOMA-based scheme, the sum throughput increases slowly or remains at an almost constant value with the increase of K. Since all the devices share the common IRS beamforming vector for data transmission in the NOMA case, the amplification coefficients of the active IRS will be restricted due to (10e) as more devices are scheduled to transmit simultaneously. To fully unleash the potential amplifying capability of the active IRS, some devices may not be scheduled or transmit at its maximum power for maximizing the sum throughput, which renders that the sum throughput increases slowly with K for the NOMA-based scheme.

Moreover, it is observed that for $K \geq 2$, the TDMA-based scheme always outperforms its NOMA-based counterpart and the performance gap becomes more pronounced with the increase of K. This is because the TDMA-based scheme can enjoy more degrees of freedom to flexibly adjust each device's dedicated IRS beamforming vector for further boosting the sum throughput compared to NOMA. As such, the TDMA-based scheme may be more preferable for the active IRS aided system at the cost of extra signalling overhead.

B. Performance Comparison for Active IRS or Passive IRS

In this subsection, we provide performance comparisons for the active IRS and passive IRS-aided energy-constrained IoT systems under different setups. For the passive IRS-aided system, we employ the algorithm proposed in [26] for obtaining the sum throughput of the TDMA-based scheme since it can characterize the upper bound performance of the passive IRS architecture in our considered scenarios. It is worth noting that the total energy consumption in active IRS aided systems is higher than that in passive IRS aided systems since the additional power is consumed at the active IRS for amplifying the signals. Therefore, it is fair to compare

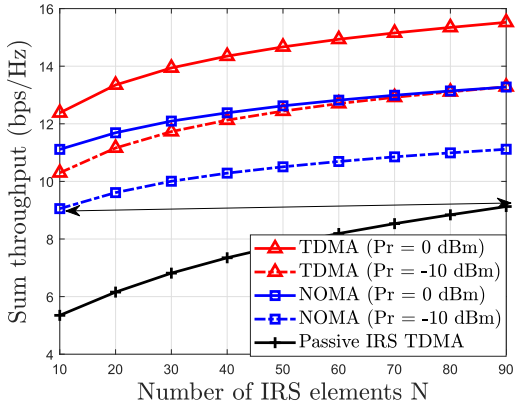


Fig. 4. Sum throughput versus the number of IRS elements with $E_k = 0.01$J and $K = 10$.

them under the same total energy budget. To this end, the additional energy consumed by the active IRS is equally added to the available energy at each device in a passive IRS aided energy-constraint IoT system, to guarantee the two systems consuming the same amount of total energy.

1) Impact of Number of IRS Elements: In Fig. 4, we plot the system sum throughput versus the number of IRS elements N for both $P_r = 0$ dBm and $P_r = -10$ dBm. It is observed that the sum throughput of all the schemes monotonically increases with N since more reflecting/amplfication elements help achieve higher IRS beamforming gains, which is beneficial for improving the power of the received signals. Moreover, one can observe that even exploiting the active IRS for the NOMA-based transmission (i.e., one active beamforming vector) can achieve significant gains over employing passive IRS for the TDMA-based transmission (i.e., K passive beamforming vectors) in terms of the system sum throughput. Considering the TDMA-based scheme relies on more IRS beamforming adjustments, the result indicates that incorporating the active IRS into IoT systems leads to a higher throughput with lower signalling overhead compared to the passive IRS. In addition, the performance gap between the two types of IRS architectures is more pronounced in the low N regime. Note that to achieve 10 bps/Hz throughput, the required number of IRS elements is reduced from 90 to 10 via replacing the passive IRS by the active IRS, which

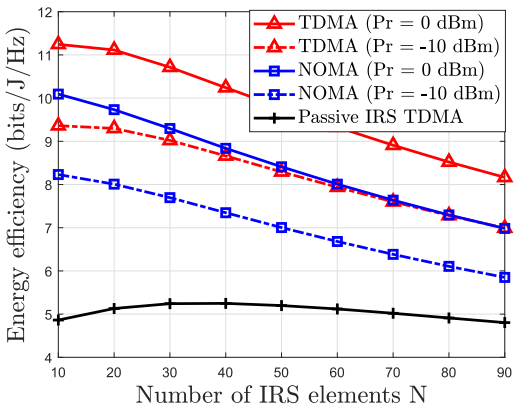


Fig. 5. Energy efficiency versus the number of IRS elements with $E_k = 0.01$J and $K = 10$.

indicates that the required number of IRS elements can be greatly reduced when employing the active IRS. The reason is that the value of the amplification coefficient at each element becomes larger when reducing the number of IRS elements. The higher amplification coefficient can effectively compensate the performance loss induced by reducing the number of IRS elements.

To further investigate the performance of active IRS aided systems, we also plot the system energy-efficiency (bits/J/Hz) versus N in Fig. 5. The energy-efficiency for both the active IRS aided TDMA and NOMA schemes in our considered setup is defined as [45]

$$\mathrm{EE}^{\mathrm{active}}_{\mathrm{TDMA/NOMA}} = \frac{R_{\mathrm{TDMA/NOMA}}}{P_r T_{\max} + \sum_{k=1}^{K} E_k + N P_I T_{\max}}, \quad (42)$$

where $P_I = 2$ mW is the circuit power required to support one IRS element [45]. By setting $P_r = 0$ in (42), the energy efficiency of the system with the conventional passive IRS can be obtained. It is observed that the energy efficiency of the active IRS still outperforms that of the conventional active IRS. The result is expected since the additional energy consumption of the active IRS over the passive IRS is actually negligible in our considered setup. Moreover, one can observe that deploying more active IRS elements would degrade the system energy efficiency since more energy needs to be consumed to support a large number of IRS elements. The result suggests that only a small number of active IRS elements is needed to guarantee the throughput performance for achieving green communications.

2) Impact of Distance Between AP and Device Center: In Fig. 6, we investigate the coverage performance of our considered active IRS-aided systems, by plotting the sum throughput versus the distance between the AP and the center of devices cluster under the Rayleigh fading and Nakagami-m fading with $m = 2$, respectively. It is observed that the sum throughput of all the considered schemes decreases as the distance increases under our considered two channel fading models. However, for both TDMA and NOMA cases, the decreasing rate of the active IRS-aided systems is much slower than that of the passive IRS-aided systems. The reason is that the power of the incident signals at each IRS element becomes smaller as the transmission distance increases, this rendering the maximum allowed amplification coefficient at each element of the active IRS to become larger. The larger amplification coefficient can effectively compensate the performance loss induced by increasing the transmission distance. Thus, the transmission coverage increases significantly by deploying the active IRS. This phenomenon demonstrates the effectiveness of the active IRS for achieving coverage extension compared to that of the conventional passive IRS.

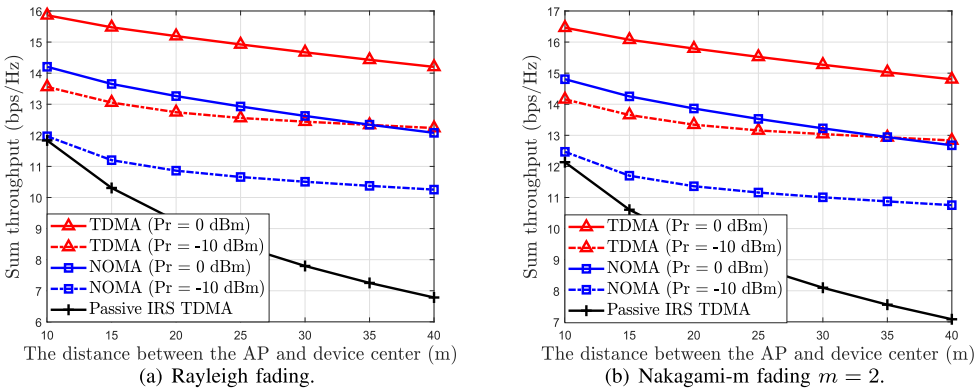
3) Impact of Available Energy E_k: In Fig. 7, we study the impact of the available energy on our considered systems, by plotting the sum throughput versus the available energy at each device. From Fig. 7, it is observed that the active IRS can significantly improve the sum throughput as compared to the case of the passive IRS, especially when the available energy at each device is low. This is because the power of the incident signal at each IRS element becomes smaller when the available energy at each device is low. As such, the maximum allowed amplification coefficient at each element of the active IRS becomes larger, which compensates the performance loss caused by the low available energy at devices. It implies that the active IRS is a more promising architecture for supporting multiple low-energy devices compared to the conventional passive IRS. In addition, the sum throughput of active IRS-aided systems for the case of $P_r = -20$ dBm is less sensitive to E_k compare to that of $P_r = 0$ dBm. The reason is that the lower value of P_r would limit the amplification amplitude at each IRS element especially when the transmit power of each device is high.

4) Impact of Deployment: In Fig. 8, the sum throughput versus the x-coordinate of the IRS x_{IRS} is plotted to demonstrate the impact of the IRS deployment on the performance of our considered systems. For conventional passive IRS-aided systems, it was shown in [6] that the IRS should be deployed in the close proximity to the AP or the device. While for the active IRS, it can be observed that the sum throughput decreases significantly when moving the IRS farther from the AP for both TDMA and NOMA cases. This is because the received signal power at the IRS gets stronger as the IRS comes closer to the devices (transmitters), which greatly limits the amplification gains at IRS elements. The main bottleneck restricting the performance becomes the deep fading of the IRS-AP link in this case. This phenomenon implies that for the active IRS, it is better to deploy it close to the AP (receiver). Additionally, the results indicate that different MA schemes have no impact on the deployment of the active IRS. Moreover, one can observe that the performance of active IRS-aided systems is even worse than that of the passive IRS if the position of the IRS is not properly selected, which highlights the importance of the active IRS deployment to unlock its full potential in maximizing the sum throughput.

C. Performance Evaluation for Hybrid MA Schemes

In this subsection, we provide numerical results for evaluating the performance of our proposed hybrid TDMA-NOMA scheme in terms of the sum throughput under different setups.

1) Evaluation of Proposed Device Grouping Scheme: Before discussing the system performance of the hybrid



(a) Rayleigh fading. (b) Nakagami-m fading $m = 2$.

Fig. 6. Sum throughput versus the distance between AP and device center with $E_k = 0.01$J and $K = 10$.

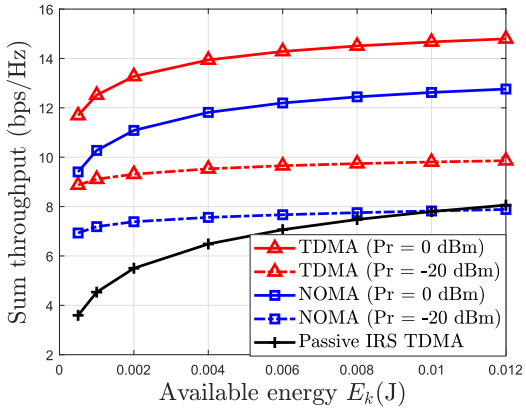


Fig. 7. Sum throughput versus the available energy E_k with $K = 10$.

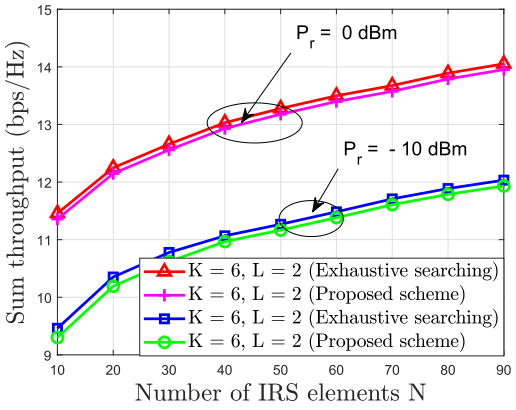


Fig. 9. Performance evaluation of proposed device under grouping scheme $E_k = 0.01$J and $K = 6$.

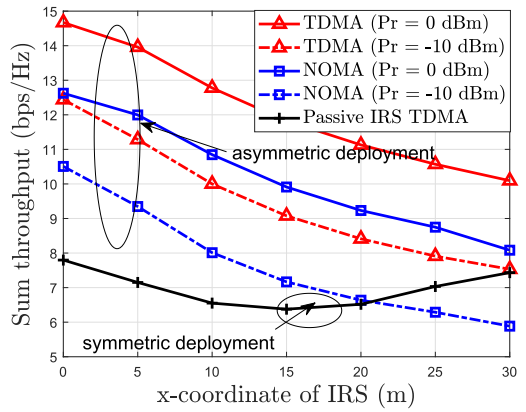


Fig. 8. Sum throughput versus the location of IRS with $E_k = 0.01$J and $K = 10$.

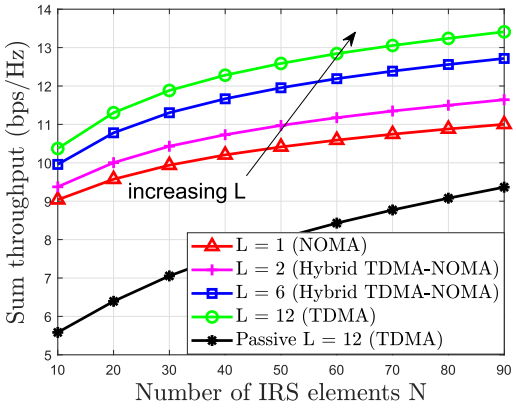


Fig. 10. Performance evaluation for hybrid MA schemes with $E_k = 0.01$J and $K = 12$.

TDMA-NOMA scheme, we first provide performance evaluation for our proposed device grouping scheme. Specifically, we compare the following two cases: 1) "Exhaustive searching" denotes the case where the optimal device grouping set is obtained through exhaustive search; 2) "Proposed scheme" denotes the the obtained results based on the proposed device grouping scheme in Section IV-B. As shown in Fig. 9, it can be observed that the performance loss of our proposed device grouping scheme is negligible compared to the results obtained by exhaustively searching, which demonstrates the effectiveness of our proposed scheme.

2) Throughput Evaluation for Hybrid MA Schemes: In Fig. 10, we evaluate the performance of the hybrid TDMA-NOMA scheme, by plotting the system sum throughput versus the number of IRS elements under the different number of device groups, i.e., L. It is observed that the sum throughput achieved by our proposed algorithm increases as L increases. The results are expected since exploiting more IRS

beamforming vectors is indeed beneficial for the throughput improvement of our considered energy-constrained IoT systems. Furthermore, one can observe that the performance gap between $L = 6$ and $L = 12$ is slight (less than 5%). Note that the number of IRS beamforming coefficients to be sent from the AP to the IRS controller is given by LN. By setting two devices in each group, the signalling overhead can be reduced by 50% compared to the TDMA-based scheme at the cost of slight performance loss, which renders the hybrid TDMA-NOMA scheme a promising approach to balance the performance-overhead tradeoff. Finally, we observe that the active IRS aided system employing NOMA significantly outperforms that of the passive IRS-aided system employing TDMA while maintaining lower signalling overhead, which highlights the potential benefits of active IRS architectures for achieving both the higher throughput and lower signaling overhead compared to the conventional passive IRS.

VI. Conclusion

Considering different MA schemes, we investigated an active IRS aided energy-constrained IoT system to maximize the sum throughput by jointly optimizing IRS beamforming vectors and resource allocation. Specifically, we first studied a couple of MA schemes, namely TDMA and NOMA. By deeply exploiting inherent properties of their associated optimization problems, we proposed two dedicated algorithms to solve them efficiently. Moreover, the theoretical performance comparison for the active IRS aided TDMA and NOMA schemes was provided. The results demonstrate that TDMA can potentially achieve a higher throughput than that of NOMA at the cost of more IRS beamforming vectors. Aiming at providing high flexibility in balancing the performance and signalling overhead tradeoff, we further develop a novel hybrid TDMA-NOMA scheme, which is applicable for any given number of IRS beamforming vectors available. The AO-based algorithm was extended to solve its associated sum throughput maximization problem. Numerical results validated our theoretical findings and unveiled the effectiveness of the active IRS architecture over the conventional passive IRS in terms of extending coverage range, reducing the requirement of reflecting elements, and supporting multiple low-energy IoT devices. Moreover, using the hybrid TDMA-NOMA scheme for assisting data transmission can be a practically appealing approach for flexibly balancing performance-overhead tradeoff, especially for IoT networks with practically large number of devices.

The results in this paper can be further extended by considering imperfect SIC for the NOMA scheme, multiple antennas at the AP, imperfect CSI with robust transmission design, and the optimization of the energy efficiency, etc, which will be left for our future work.

Appendix A: Proof of Lemma 1

We first prove that all devices will be scheduled at the optimal solution, i.e., $\tau_k^* > 0, \forall k$. We show it by contradiction. Suppose that $\Xi^* = \{\tau_k^*, p_k^*, \mathbf{v}_k^*\}$ achieves the optimal solution of problem (11) and there exists a device j who will not be scheduled, i.e., $\tau_j^* = 0$. Then, we construct a different solution $\tilde{\Xi} = \{\tilde{\tau}_k, \tilde{p}_k, \tilde{\mathbf{v}}_k\}$, where $\tilde{\mathbf{v}}_k = \mathbf{v}_k^* \, (k \neq j)$, $\tilde{\mathbf{v}}_j = \mathbf{0}$, and

$$\tilde{\tau}_k = \begin{cases} \tau_k^*, & k \neq j, k \neq m, k \in \mathcal{K}, \\ \tau_k^* - \Delta\tau_k, & k = m, m \in \mathcal{K}, \\ \Delta\tau_k, & k = j, j \in \mathcal{K}, \end{cases} \tag{43}$$

$$\tilde{p}_k = \begin{cases} p_k^*, & k \neq j, k \in \mathcal{K}, \\ \frac{E_j}{\Delta\tau_m}, & k = j, j \in \mathcal{K}, m \in \mathcal{K}. \end{cases} \tag{44}$$

It can be readily verified that the newly constructed solution $\tilde{\Xi}$ is also a feasible solution for problem (11) since it satisfies all the constraints therein. Since the solutions regarding the transmit power, the time allocation, and beamforming vector for any device $k \neq j, k \neq m$ remain unchanged in Ξ^* and $\tilde{\Xi}$, we only need to compare the throughput contributed by device j and device m for the corresponding two solutions. For the solution Ξ^*, we have

$$R_m^* + R_j^* = (\tau_m^* - \Delta\tau_m) \log_2 \left(1 + \frac{p_m \left|h_{d,k} + \mathbf{v}_m^H \mathbf{q}_m\right|^2}{\sigma^2 + \sigma_r^2 \mathbf{v}_m^H \mathbf{G} \mathbf{v}_m}\right) + \Delta\tau_m \log_2 \left(1 + \frac{p_m \left|h_{d,k} + \mathbf{v}_m^H \mathbf{q}_m\right|^2}{\sigma^2 + \sigma_r^2 \mathbf{v}_m^H \mathbf{G} \mathbf{v}_m}\right). \tag{45}$$

For the newly constructed solution $\tilde{\Xi}$, we have

$$\tilde{R}_m + \tilde{R}_j = (\tau_m^* - \Delta\tau_m) \log_2 \left(1 + \frac{p_m \left|h_{d,k} + \mathbf{v}_m^H \mathbf{q}_m\right|^2}{\sigma^2 + \sigma_r^2 \mathbf{v}_m^H \mathbf{G} \mathbf{v}_m}\right) + \Delta\tau_m \log_2 \left(1 + \frac{E_j}{\Delta\tau_m} |h_{d,j}|^2\right). \tag{46}$$

For a given positive value

$$\varepsilon = \frac{E_j \left(\sigma^2 + \sigma_r^2 \mathbf{v}_m^H \mathbf{G} \mathbf{v}_m\right) |h_{d,j}|^2}{p_m |h_{d,k} + \mathbf{v}_m^H \mathbf{q}_m|^2}, \tag{47}$$

it can be easily verified that $\tilde{R}_m + \tilde{R}_j > R_m^* + R_j^*$ when $\Delta\tau_m < \varepsilon$. This means that the newly constructed solution $\tilde{\Xi}$ achieves a higher throughput than Ξ^*, which contradicts the assumption that Ξ^* is the optimal solution. As such, we have $\tau_k^* > 0, \forall k$.

We next prove $p_k^* = E_k / \tau_k^*$ by contradiction. Suppose that $\{p_k^*, \mathbf{v}_k^*\}$ is the optimal transmit power and IRS beamforming vector for device k and $p_k^* < E_k / \tau_k^*$. We can always construct a new solution denoted by $\{\tilde{p}_k, \tilde{\mathbf{v}}_k\}$ which satisfies $p_k^* < \tilde{p}_k \leq E_k / \tau_k^*$ and $\tilde{\mathbf{v}}_k = \sqrt{p_k^* / \tilde{p}_k} \mathbf{v}_k^*$. It can be readily verified that

$$\begin{aligned} &\tilde{p}_k \tilde{\mathbf{v}}_k^H \mathbf{H}_{r,k} \tilde{\mathbf{v}}_k + \sigma_r^2 \|\tilde{\mathbf{v}}_k\|^2 \\ &\quad < p_k^* \mathbf{v}_k^{*H} \mathbf{H}_{r,k} \mathbf{v}_k^* + \sigma_r^2 \|\mathbf{v}_k^*\|^2 \leq P_r, \end{aligned} \tag{48}$$

which indicates that $\{\tilde{p}_k, \tilde{\mathbf{v}}_k\}$ is feasible for problem (11). We further compare the objective values for the two solutions, namely $\{p_k^*, \mathbf{v}_k^*\}$ and $\{\tilde{p}_k, \tilde{\mathbf{v}}_k\}$, as follows

$$\frac{p_k^* \left|h_{d,k} + \mathbf{v}_k^{*H} \mathbf{q}_k\right|^2}{\sigma^2 + \sigma_r^2 \mathbf{v}_k^{*H} \mathbf{G} \mathbf{v}_k^*} \overset{(a)}{<} \frac{\tilde{p}_k \left|h_{d,k} + \tilde{\mathbf{v}}_k^H \mathbf{q}_k\right|^2}{\sigma^2 + \sigma_r^2 \tilde{\mathbf{v}}_k^H \mathbf{G} \tilde{\mathbf{v}}_k}, \tag{49}$$

where inequality (a) holds due to $p_k^* |h_{d,k}|^2 < \tilde{p}_k |h_{d,k}|^2$, $p_k^* \mathrm{Re}\left(h_{d,k} \mathbf{v}_k^{*H} \mathbf{q}_k\right) < \tilde{p}_k \mathrm{Re}\left(h_{d,k} \tilde{\mathbf{v}}_k^H \mathbf{q}_k\right)$, $p_k^* \left|\mathbf{v}_k^{*H} \mathbf{q}_k\right|^2 = \tilde{p}_k \left|\tilde{\mathbf{v}}_k^H \mathbf{q}_k\right|^2$, and $\mathbf{v}_k^{*H} \mathbf{G} \mathbf{v}_k^* > \tilde{\mathbf{v}}_k^H \mathbf{G} \tilde{\mathbf{v}}_k$. This means that the

newly constructed solution $\{\tilde{p}_k, \tilde{\mathbf{v}}_k\}$ achieves a higher objective value than that of $\{p_k^*, \mathbf{v}_k^*\}$, which contradicts the assumption that $\{p_k^*, \mathbf{v}_k^*\}$ is optimal. Thus, the optimal transmit power satisfies $p_k^* = E_k/\tau_k^*, \forall k$.

Appendix B: Proof of Proposition 1

To prove Proposition 1, we first introduce a pair of problem formulations corresponding to problem (18) and (32) without constraints (18c) and (32c) as follows, respectively,

$$\max_{\tau, p_k, \mathbf{v}} \quad \tau \log_2\left(1 + \frac{\sum_{k=1}^K p_k \left|h_{d,k} + \mathbf{v}^H \mathbf{q}_k\right|^2}{\sigma^2 + \sigma_r^2 \mathbf{v}^H \mathbf{G} \mathbf{v}}\right) \tag{50a}$$

$$\text{s.t.} \quad (10b), (10c), (10d). \tag{50b}$$

$$\max_{\tau_k, p_k, \mathbf{v}} \quad \sum_{k=1}^K \tau_k \log_2\left(1 + \frac{p_k \left|h_{d,k} + \mathbf{v}^H \mathbf{q}_k\right|^2}{\sigma^2 + \sigma_r^2 \mathbf{v}^H \mathbf{G} \mathbf{v}}\right) \tag{51a}$$

$$\text{s.t.} \quad (9b), (9c), (9d). \tag{51b}$$

The optimal values of problem (50) and (51) are denoted by $R_{\text{NOMA}}^{\text{ub}}$ and $R_{\text{TDMA}}^{\text{ub}}$, respectively. Firstly, we can prove $R_{\text{TDMA}}^{\text{ub}} = R_{\text{NOMA}}^{\text{ub}}$ by showing $R_{\text{TDMA}}^{\text{ub}} \le R_{\text{NOMA}}^{\text{ub}}$ and $R_{\text{TDMA}}^{\text{ub}} \ge R_{\text{NOMA}}^{\text{ub}}$.

The procedure starts by showing that $R_{\text{TDMA}}^{\text{ub}} \le R_{\text{NOMA}}^{\text{ub}}$. For problem (51), the optimal transmit power of each device, denoted by $\breve{p}_k$, can be expressed as $\breve{p}_k = E_k/\tau_k$, because each device will deplete all of its energy. To this end, we discuss some properties about the optimal time allocation for problem (51), i.e., $\breve{\tau}_k$. Given the optimal IRS beamforming vector $\mathbf{v} = \breve{\mathbf{v}}$, problem (51) can be simplified by optimizing τ_k as follows

$$\max_{\tau_k} \quad \sum_{k=1}^K \tau_k \log_2\left(1 + \frac{E_k \left|h_{d,k} + \breve{\mathbf{v}}^H \mathbf{q}_k\right|^2}{\tau_k \left(\sigma^2 + \sigma_r^2 \breve{\mathbf{v}}^H \mathbf{G} \breve{\mathbf{v}}\right)}\right) \tag{52a}$$

$$\text{s.t.} \quad (9c). \tag{52b}$$

Note that problem (52) is a convex optimization problem and its Lagrangian function is

$$\mathcal{L}(\tau_k, \lambda) = \sum_{k=1}^K \tau_k \log_2\left(1 + \frac{E_k \left|h_{d,k} + \breve{\mathbf{v}}^H \mathbf{q}_k\right|^2}{\tau_k \left(\sigma^2 + \sigma_r^2 \breve{\mathbf{v}}^H \mathbf{G} \breve{\mathbf{v}}\right)}\right) + \lambda\left(T_{\max} - \sum_{k=1}^K \tau_k\right), \tag{53}$$

where $\lambda \ge 0$ is the dual variable associated with (52b). According to Karush-Kuhn-Tucker (KKT) conditions, we have

$$\frac{\partial \mathcal{L}(\tau_k, \lambda)}{\partial \tau_k} = \Gamma(\Upsilon_k) \triangleq \log_2(1 + \Upsilon_k) - \frac{\Upsilon_k}{(1 + \Upsilon_k)\ln 2} - \lambda = 0, \tag{54}$$

where

$$\Upsilon_k = \frac{E_k \left|h_{d,k} + \breve{\mathbf{v}}^H \mathbf{q}_k\right|^2}{\tau_k \left(\sigma^2 + \sigma_r^2 \breve{\mathbf{v}}^H \mathbf{G} \breve{\mathbf{v}}\right)}, \tag{55}$$

can be regarded as the received signal-to-noise ratio (SNR) of device k. Since $\Gamma(\Upsilon_k)$ is an increasing function with respect to Υ_k and $\Gamma(0) \le 0$, equation $\Gamma(\Upsilon_k) = 0$ has a unique solution, which implies that all devices share the same SNR at the optimal solution, i.e.,

$$\frac{E_k \left|h_{d,k} + \breve{\mathbf{v}}^H \mathbf{q}_k\right|^2}{\breve{\tau}_k \left(\sigma^2 + \sigma_r^2 \breve{\mathbf{v}}^H \mathbf{G} \breve{\mathbf{v}}\right)} = \frac{E_j \left|h_{d,k} + \breve{\mathbf{v}}^H \mathbf{q}_j\right|^2}{\breve{\tau}_j \left(\sigma^2 + \sigma_r^2 \breve{\mathbf{v}}^H \mathbf{G} \breve{\mathbf{v}}\right)}, \forall k, j \in \mathcal{K}. \tag{56}$$

As such, the optimal value of problem (51) can be written as

$$R_{\text{TDMA}}^{\text{ub}} = \breve{\tau} \log_2\left(1 + \frac{\sum_{k=1}^K E_k \left|h_{d,k} + \breve{\mathbf{v}}^H \mathbf{q}_k\right|^2}{\breve{\tau}\left(\left(\sigma^2 + \sigma_r^2 \breve{\mathbf{v}}^H \mathbf{G} \breve{\mathbf{v}}\right)\right)}\right), \tag{57}$$

where $\breve{\tau} = \sum_{k=1}^K \breve{\tau}_k$. It can be verified that $\left\{\tau = \breve{\tau}, p_k = E_k \Big/ \breve{\tau}, \mathbf{v} = \breve{\mathbf{v}}\right\}$ is one feasible solution for problem (50), which yields $R_{\text{TDMA}}^{\text{ub}} \le R_{\text{NOMA}}^{\text{ub}}$.

Next, we show that $R_{\text{TDMA}}^{\text{ub}} \ge R_{\text{NOMA}}^{\text{ub}}$. For problem (50), it can be easily shown that each device will deplete all of its energy and thus the optimal solution of problem (50) is denoted by $\left\{\tau = \widehat{\tau}, p_k = E_k \Big/ \widehat{\tau}, \mathbf{v} = \widehat{\mathbf{v}}\right\}$. Based on the optimal solution of problem (50), we can always construct a new solution, which satisfies $\widehat{\tau} = \sum_{k=1}^K \widehat{\tau}_k$ and $\widehat{p}_k = E_k \Big/ \widehat{\tau}_k$, so that all devices share the same received SNR in the TDMA scheme, i.e.,

$$\frac{E_k \left|h_{d,k} + \widehat{\mathbf{v}}^H \mathbf{q}_k\right|^2}{\widehat{\tau}_k \left(\sigma^2 + \sigma_r^2 \widehat{\mathbf{v}}^H \mathbf{G} \widehat{\mathbf{v}}\right)} = \frac{E_j \left|h_{d,k} + \widehat{\mathbf{v}}^H \mathbf{q}_j\right|^2}{\widehat{\tau}_j \left(\sigma^2 + \sigma_r^2 \widehat{\mathbf{v}}^H \mathbf{G} \widehat{\mathbf{v}}\right)}, \forall k, j \in \mathcal{K}. \tag{58}$$

Note that $\left\{\tau_k = \widehat{\tau}_k, p_k = E_k \Big/ \widehat{\tau}_k, \mathbf{v} = \widehat{\mathbf{v}}\right\}$ is also feasible for problem (51) and always exists, which yields

$$R_{\text{NOMA}}^{\text{ub}} = \sum_{k=1}^K \widehat{\tau}_k \log_2\left(1 + \frac{E_k \left|h_{d,k} + \widehat{\mathbf{v}}^H \mathbf{q}_k\right|^2}{\widehat{\tau}_k \left(\sigma^2 + \sigma_r^2 \widehat{\mathbf{v}}^H \mathbf{G} \widehat{\mathbf{v}}\right)}\right). \tag{59}$$

Thus, at the optimal solution of problem (51), it follows that $R_{\text{TDMA}}^{\text{ub}} \ge R_{\text{NOMA}}^{\text{ub}}$.

Given $R_{\text{TDMA}}^{\text{ub}} \le R_{\text{NOMA}}^{\text{ub}}$ and $R_{\text{TDMA}}^{\text{ub}} \ge R_{\text{NOMA}}^{\text{ub}}$, we have $R_{\text{TDMA}}^{\text{ub}} = R_{\text{NOMA}}^{\text{ub}}$. Regarding their optimal solutions, we have $\widehat{\tau} = \sum_{k=1}^K \breve{\tau}_k$, $\widehat{\mathbf{v}} = \breve{\mathbf{v}}$, and $\widehat{e}_k = \breve{e}_k$, where $\widehat{e}_k = \widehat{\tau}\widehat{p}_k$ and $\breve{e}_k = \breve{\tau}_k \breve{p}_k$. Then, we add constraints (18c) and (32c) in problems (50) and (51), respectively, which correspondingly results in problems (18) and (32). By letting $e_k = \tau p_k$ in problem (18) and $e_k = \tau_k p_k$ in problem (32), constraints (18c) and (32c) can be equivalently rewritten as

$$\sum_{k=1}^K e_k \mathbf{v}^H \mathbf{H}_{r,k} \mathbf{v} + \tau \sigma_r^2 \|\mathbf{v}\|^2 \le P_r \tau, \tag{60a}$$

$$e_k \mathbf{v}^H \mathbf{H}_{r,k} \mathbf{v} + \tau_k \sigma_r^2 \|\mathbf{v}\|^2 \le P_r \tau_k, \forall k \in \mathcal{K}. \tag{60b}$$

For any feasible solution $\{\tilde{\tau}_k, \tilde{e}_k, \tilde{\mathbf{v}}\}$ for problem (32), we have

$$\sum_{k=1}^K \tilde{\tau}_k \log_2\left(1 + \frac{\tilde{e}_k \left|h_{d,k} + \mathbf{v}^H \mathbf{q}_k\right|^2}{\tilde{\tau}_k \left(\sigma^2 + \sigma_r^2 \mathbf{v}^H \mathbf{G} \mathbf{v}\right)}\right)$$

$$\overset{(a)}{\leq} \sum_{k=1}^{K} \tilde{\tau}_k \log_2 \left(1 + \frac{\sum_{k=1}^{K} \tilde{e}_k \left|h_{d,k} + \mathbf{v}^H \mathbf{q}_k\right|^2}{\sum_{k=1}^{K} \tilde{\tau}_k \left(\sigma^2 + \sigma_r^2 \mathbf{v}^H \mathbf{G} \mathbf{v}\right)}\right)$$
$$= \tilde{\tau} \log_2 \left(1 + \frac{\sum_{k=1}^{K} \tilde{e}_k \left|h_{d,k} + \mathbf{v}^H \mathbf{q}_k\right|^2}{\tilde{\tau} \left(\sigma^2 + \sigma_r^2 \mathbf{v}^H \mathbf{G} \mathbf{v}\right)}\right), \quad (61)$$

where (a) follows from the inequality of arithmetic and geometric means, and $\tilde{\tau} = \sum_{k=1}^{K} \tilde{\tau}_k$. We have $\sum_{k=1}^{K} \tilde{e}_k \tilde{\mathbf{v}}^H \mathbf{H}_{r,k} \tilde{\mathbf{v}} + \tilde{\tau} \sigma_r^2 \|\tilde{\mathbf{v}}\|^2 \leq P_r \tilde{\tau}$ by taking the summation of all $\forall k \in \mathcal{K}$ in (60b), which implies that $\{\tilde{\tau}, \tilde{e}_k, \tilde{\mathbf{v}}\}$ is guaranteed as a feasible solution for (18). As such, the optimal value of problem (18) is no smaller than that of (32), i.e., $R^*_{\mathrm{NOMA}} \geq R^{(\mathrm{lb})*}_{\mathrm{TDMA}}$. If

$$\frac{\tilde{e}_k \left|h_{d,k} + \mathbf{v}^H \mathbf{q}_k\right|^2}{\tilde{\tau}_k \left(\sigma^2 + \sigma_r^2 \mathbf{v}^H \mathbf{G} \mathbf{v}\right)} \neq \frac{\tilde{e}_j \left|h_{d,j} + \mathbf{v}^H \mathbf{q}_j\right|^2}{\tilde{\tau}_j \left(\sigma^2 + \sigma_r^2 \mathbf{v}^H \mathbf{G} \mathbf{v}\right)}, \exists k, j \in \mathcal{K}, \quad (62)$$

we have $R^*_{\mathrm{NOMA}} > R^{(\mathrm{lb})*}_{\mathrm{TDMA}}$, i.e., the optimal value of (18) is strictly larger than that of (32).

When $\left\{\breve{\tau}_k, \breve{e}_k, \breve{\mathbf{v}}\right\}$ satisfies constraint (60b), it is also an optimal solution for problem (18) and thus $R^{(\mathrm{lb})*}_{\mathrm{TDMA}} = R^{\mathrm{ub}}_{\mathrm{TDMA}}$. In this case, $\widehat{\tau} = \sum_{k=1}^{K} \breve{\tau}_k$, $\widehat{\mathbf{v}} = \breve{\mathbf{v}}$, and $\widehat{e}_k = \breve{e}_k$ also satisfy all constraints in problem (18), which indicates that $R^{(\mathrm{lb})*}_{\mathrm{TDMA}} = R^{\mathrm{ub}}_{\mathrm{TDMA}} = R^{\mathrm{ub}}_{\mathrm{NOMA}} = R^*_{\mathrm{NOMA}}$. Therefore, (35) is a sufficient condition for $R^{(\mathrm{lb})*}_{\mathrm{TDMA}} = R^*_{\mathrm{NOMA}}$, which thus completes the proof.

REFERENCES

[1] L. Chettri and R. Bera, "A comprehensive survey on Internet of Things (IoT) toward 5G wireless systems," *IEEE Internet Things J.*, vol. 7, no. 1, pp. 16–32, Jan. 2020.

[2] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, "A survey on Internet of Things: Architecture, enabling technologies, security and privacy, and applications," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1125–1142, Apr. 2017.

[3] Q. Wu, W. Chen, D. W. K. Ng, J. Li, and R. Schober, "User-centric energy efficiency maximization for wireless powered communications," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6898–6912, Oct. 2016.

[4] K. Chi, Z. Chen, K. Zheng, Y.-H. Zhu, and J. Liu, "Energy provision minimization in wireless powered communication networks with network throughput demand: TDMA or NOMA?" *IEEE Trans. Commun.*, vol. 67, no. 9, pp. 6401–6414, Sep. 2019.

[5] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Jan. 2020.

[6] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, May 2021.

[7] Y. Liu *et al.*, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1546–1577, 3rd Quart., 2021.

[8] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.

[9] Q. Q. Wu and R. Zhang, "Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1838–1851, May 2020.

[10] C. Pan *et al.*, "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5218–5233, Aug. 2020.

[11] M. Hua, Q. Wu, D. W. K. Ng, J. Zhao, and L. Yang, "Intelligent reflecting surface-aided joint processing coordinated multipoint transmission," *IEEE Trans. Commun.*, vol. 69, no. 3, pp. 1650–1665, Mar. 2021.

[12] H. Xie, J. Xu, and Y.-F. Liu, "Max-min fairness in IRS-aided multi-cell MISO systems with joint transmit and reflective beamforming," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 1379–1393, Feb. 2020.

[13] X. Yu, D. Xu, Y. Sun, D. W. K. Ng, and R. Schober, "Robust and secure wireless communications via intelligent reflecting surfaces," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2637–2652, Nov. 2020.

[14] X. Guan, Q. Wu, and R. Zhang, "Intelligent reflecting surface assisted secrecy communication: Is artificial noise helpful or not?" *IEEE Wireless Commun. Lett.*, vol. 9, no. 6, pp. 778–782, Jun. 2020.

[15] S. Hu, Z. Wei, Y. Cai, C. Liu, D. Wing Kwan Ng, and J. Yuan, "Robust and secure sum-rate maximization for multiuser MISO downlink systems with self-sustainable IRS," 2021, *arXiv:2101.10549*.

[16] X. Lu, W. Yang, X. Guan, Q. Wu, and Y. Cai, "Robust and secure beamforming for intelligent reflecting surface aided mmWave MISO systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 12, pp. 2068–2072, Sep. 2020.

[17] W. Wang and W. Zhang, "Joint beam training and positioning for intelligent reflecting surfaces assisted millimeter wave communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6282–6297, Oct. 2021.

[18] Z. Wei *et al.*, "Sum-rate maximization for IRS-assisted UAV OFDMA communication systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 4, pp. 2530–2550, Apr. 2020.

[19] X. Mu, Y. Liu, L. Guo, J. Lin, and H. V. Poor, "Intelligent reflecting surface enhanced multi-UAV NOMA networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 10, pp. 3051–3066, Oct. 2021.

[20] Q. Wu, X. Guan, and R. Zhang, "Intelligent reflecting surface-aided wireless energy and information transmission: An overview," *Proc. IEEE*, vol. 110, no. 1, pp. 150–170, Jan. 2022.

[21] Q. Wu and R. Zhang, "Weighted sum power maximization for intelligent reflecting surface aided SWIPT," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 586–590, Oct. 2020.

[22] S. Zargari, A. Khalili, and R. Zhang, "Energy efficiency maximization via joint active and passive beamforming design for multiuser MISO IRS-aided SWIPT," *IEEE Wireless Commun. Lett.*, vol. 10, no. 3, pp. 557–561, Jun. 2021.

[23] Q. Wu and R. Zhang, "Joint active and passive beamforming optimization for intelligent reflecting surface assisted SWIPT under QoS constraints," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1735–1748, Aug. 2020.

[24] B. Lyu, P. Ramezani, D. T. Hoang, S. Gong, Z. Yang, and A. Jamalipour, "Optimized energy and information relaying in self-sustainable IRS-empowered WPCN," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 619–633, Jan. 2021.

[25] Y. Zheng, S. Bi, Y.-J.-A. Zhang, X. Lin, and H. Wang, "Joint beamforming and power control for throughput maximization in IRS-assisted MISO WPCNs," *IEEE Internet Things J.*, vol. 8, no. 10, pp. 8399–8410, Aug. 2021.

[26] Q. Wu, X. Zhou, W. Chen, J. Li, and X. Zhang, "IRS-aided WPCNs: A new optimization framework for dynamic IRS beamforming," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 4725–4739, Jul. 2022.

[27] M. Hua and Q. Wu, "Joint dynamic passive beamforming and resource allocation for IRS-aided full-duplex WPCN," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 4829–4843, Jul. 2022.

[28] X. Hu, C. Masouros, and K.-K. Wong, "Reconfigurable intelligent surface aided mobile edge computing: From optimization-based to location-only learning-based solutions," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 3709–3725, Jun. 2021.

[29] Z. Chu, P. Xiao, M. Shojafar, D. Mi, J. Mao, and W. Hao, "Intelligent reflecting surface assisted mobile edge computing for Internet of Things," *IEEE Wireless Commun. Lett.*, vol. 10, no. 3, pp. 619–623, Jun. 2021.

[30] G. Chen, Q. Wu, W. Chen, D. W. K. Ng, and L. Hanzo, "IRS-aided wireless powered MEC systems: TDMA or NOMA for computation offloading?" *IEEE Trans. Wireless Commun.*, early access.

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

[32] Z. Zhang *et al.*, "Active RIS vs. passive RIS: Which will prevail in 6G?" 2021, *arXiv:2103.15154*.

[33] C. You and R. Zhang, "Wireless communication aided by intelligent reflecting surface: Active or passive?" 2021, *arXiv:2106.10963*.

[34] 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 Commun. Lett.*, vol. 26, no. 5, pp. 1150–1154, May 2022.

[35] Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Nonorthogonal multiple access for 5G and beyond," *Proc. IEEE*, vol. 105, no. 12, pp. 2347–2381, Dec. 2017.
[36] F. Wang, J. Xu, and Z. Ding, "Multi-antenna NOMA for computation offloading in multiuser mobile edge computing systems," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 2450–2463, Mar. 2018.
[37] F. Fang, Y. Xu, Z. Ding, C. Shen, M. Peng, and G. K. Karagiannidis, "Optimal resource allocation for delay minimization in NOMA-MEC networks," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7867–7881, Dec. 2020.
[38] L. Shi, Y. Ye, X. Chu, and G. Lu, "Computation energy efficiency maximization for a NOMA-based WPT-MEC network," *IEEE Internet Things J.*, vol. 8, no. 13, pp. 10731–10744, Aug. 2021.
[39] R. Duan, J. Wang, C. Jiang, H. Yao, Y. Ren, and Y. Qian, "Resource allocation for multi-UAV aided IoT NOMA uplink transmission systems," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 7025–7037, Jun. 2019.
[40] X. Mu, Y. Liu, L. Guo, J. Lin, and Z. Ding, "Energy-constrained UAV data collection systems: NOMA and OMA," *IEEE Trans. Veh. Technol.*, vol. 70, no. 7, pp. 6898–6912, Jul. 2021.
[41] L. Wei, C. Huang, G. C. Alexandropoulos, C. Yuen, Z. Zhang, and M. Debbah, "Channel estimation for RIS-empowered multi-user MISO wireless communications," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 4144–4157, Jun. 2021.
[42] M. Mohseni, R. Zhang, and J. M. Cioffi, "Optimized transmission for fading multiple-access and broadcast channels with multiple antennas," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1627–1639, Aug. 2006.
[43] Y. Huang and D. P. Palomar, "Rank-constrained separable semidefinite programming with applications to optimal beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 664–678, Feb. 2010.
[44] M. A. Sedaghat and R. R. Müller, "On user pairing in uplink NOMA," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3474–3486, May 2018.
[45] D. Xu, X. Yu, D. Wing Kwan Ng, and R. Schober, "Resource allocation for active IRS-assisted multiuser communication systems," 2021, *arXiv:2108.13033*.

Guangji Chen received the Ph.D. degree in information and communication engineering from the University of Science and Technology of China (USTC), Hefei, China, in 2020.

From 2020 to 2021, he was a Senior Engineer at Huawei Technologies Company Ltd., Nanjing, China. He is currently a Post-Doctoral Researcher with the State Key Laboratory of Internet of Things for Smart City, University of Macau, Macau, China. His current research interests include intelligent reflecting surface (IRS), non-orthogonal multiple access, ultra dense network (UDN), and stochastic geometry.

Qingqing Wu (Senior Member, IEEE) received the B.Eng. degree in electronic engineering from the South China University of Technology in 2012 and the Ph.D. degree in electronic engineering from Shanghai Jiao Tong University (SJTU) in 2016. From 2016 to 2020, he was a Research Fellow of the Department of Electrical and Computer Engineering, National University of Singapore. He is currently an Assistant Professor with the State Key Laboratory of Internet of Things for Smart City, University of Macau. He has coauthored more than 100 IEEE journal articles with 25 ESI highly cited papers and eight ESI hot papers, which have received more than 15,000 Google citations. His current research interests include intelligent reflecting surface (IRS), unmanned aerial vehicle (UAV) communications, and MIMO transceiver design. He was listed as the Clarivate ESI Highly Cited Researcher in 2021 and 2022, the Most Influential Scholar Award in AI-2000 by Aminer in 2021, and Worlds Top 2% Scientist by Stanford University in 2020 and 2021.

He was a recipient of the IEEE Communications Society Young Author Best Paper Award in 2021, the Outstanding Ph.D. Thesis Award of China Institute of Communications in 2017, the Outstanding Ph.D. Thesis Funding in SJTU in 2016, the IEEE ICCC Best Paper Award in 2021, and IEEE WCSP Best Paper Award in 2015. He was the Exemplary Editor of IEEE COMMUNICATIONS LETTERS in 2019 and the Exemplary Reviewer of several IEEE journals. He serves as an Associate Editor for IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE COMMUNICATIONS LETTERS, IEEE WIRELESS COMMUNICATIONS LETTERS, IEEE OPEN JOURNAL OF COMMUNICATIONS SOCIETY (OJ-COMS), and IEEE OPEN JOURNAL OF VEHICULAR TECHNOLOGY (OJVT). He is the Lead Guest Editor for IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS on "UAV Communications in 5G and Beyond Networks," and the Guest Editor for IEEE OPEN JOURNAL OF VEHICULAR TECHNOLOGY on 6G Intelligent Communication and IEEE OPEN JOURNAL OF COMMUNICATIONS SOCIETY on Reconfigurable Intelligent Surface-Based Communications for 6G Wireless Networks. He is the Workshop Co-Chair of IEEE ICC 2019-2022 Workshop on Integrating UAVs into 5G and Beyond and the Workshop Co-Chair of IEEE GLOBECOM 2020 and ICC 2021 Workshop on Reconfigurable Intelligent Surfaces for Wireless Communication for Beyond 5G. He serves as the Workshops and the Symposia Officer of Reconfigurable Intelligent Surfaces Emerging Technology Initiative and the Research Blog Officer of Aerial Communications Emerging Technology Initiative. He is the IEEE Communications Society Young Professional Chair in Asia Pacific Region.

Chong He (Member, IEEE) received the B.Sc. degree in electronic and information engineering and the M.S. degree in electromagnetic and microwave technology from the Huazhong University of Science and Technology, Wuhan, China, in 2007 and 2009, respectively, and the Ph.D. degree in electromagnetic and microwave technology from Shanghai Jiao Tong University, Shanghai, China, in 2015. From 2016 to 2018, he was a Post-Doctoral Researcher with the Department of Electronic Engineering, Shanghai Jiao Tong University, where he has been joining the School of Electronic Information and Electrical Engineering as an Assistant Professor since 2019. He has authored or coauthored over 80 articles in IEEE journals and conferences. His research interests include phased arrays, reconfigurable intelligent surface, DOA estimation, wireless location, and multiple access wireless communications.

He has served as a Reviewer of IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, IEEE ANTENNAS AND WIRELESS PROPAGATION LETTERS, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and IEEE WIRELESS COMMUNICATIONS LETTERS. He served as a Guest Editor in a special issue of IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (JSAC) in 2021 and a TPC Member of IEEE GC 2020 MWN Symposium and 2021 IEEE Wireless Communications and Networking Conference.

Wen Chen (Senior Member, IEEE) is currently a tenured Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University, China, where he is also the Director of the Broadband Access Network Laboratory. He has published more than 110 papers in IEEE journals and more than 120 papers in IEEE conferences, with citations more than 8000 in Google scholar. His research interests include multiple access, wireless AI, and metasurface communications. He is a fellow of the Chinese Institute of Electronics and the Distinguished Lecturer of IEEE Communications Society and IEEE Vehicular Technology Society. He is also the Shanghai Chapter Chair of IEEE Vehicular Technology Society, an Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE ACCESS, and IEEE OPEN JOURNAL OF VEHICULAR TECHNOLOGY.

Jie Tang (Senior Member, IEEE) received the B.Eng. degree from the South China University of Technology, China, the M.Sc. degree from the University of Bristol, U.K., and the Ph.D. degree from Loughborough University, U.K.

From 2013 to 2015, he was a Research Associate with the School of Electrical and Electronic Engineering, University of Manchester, U.K. He is currently a Professor with the School of Electronic and Information Engineering, South China University of Technology. His current research interests include SWIPT, UAV communications, NOMA, and reconfigurable intelligent surface. He is also serving as an Editor for IEEE WIRELESS COMMUNICATIONS LETTERS, IEEE SYSTEMS JOURNAL, and *EURASIP Journal on Wireless Communications and Networking*. He is the Guest Editor of IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING and IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY. He also served as the Track Co-Chair of IEEE VTC-Spring 2018 and 2022, the Symposium Co-Chair of IEEE/CIC ICCC 2020, IEEE ComComAp 2019, the TPC Co-Chair of EAI GreeNets 2019, and the Workshop Co-Chair of IEEE ICCC/CIC 2019. He received the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2021. Also, he was a co-recipient of best paper awards at the ICNC 2018, CSPS 2018, WCSP 2019, 6GN 2020, and AICON2021.

Shi Jin (Senior Member, IEEE) received the B.S. degree in communications engineering from the Guilin University of Electronic Technology, Guilin, China, in 1996, the M.S. degree from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2003, and the Ph.D. degree in information and communications engineering from Southeast University, Nanjing, in 2007. From June 2007 to October 2009, he was a Research Fellow of the Adastral Park Research Campus, University College London, London, U.K. He is currently a Faculty Member of the National Mobile Communications Research Laboratory, Southeast University. His research interests include space time wireless communications, random matrix theory, and information theory. He serves as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE COMMUNICATIONS LETTERS, and *IET Communications*. He and his coauthors have been awarded the 2011 IEEE Communications Society Stephen O. Rice Prize Paper Award in the field of communication theory and the 2010 Young Author Best Paper Award by the IEEE Signal Processing Society.