IEEE Open Journal of the Communications Society

Received XX Month, XXXX; revised XX Month, XXXX; accepted XX Month, XXXX; Date of publication XX Month, XXXX; date of current version XX Month, XXXX.

Digital Object Identifier 10.1109/OJCOMS.2022.1234567

Throughput Maximisation in RIS-assisted NOMA-THz Communication Network

Tan Do-Duy[1], **Antonino Masaracchia**[2, 3] ***(Senior Member IEEE)***, **Berk Canberk**[4, 5] ***(Senior Member IEEE)***, **Long D. Nguyen**[6] ***(Member IEEE)***, **and Trung Q. Duong**[2, 7] ***(Fellow IEEE)***

[1]Ho Chi Minh City University of Technology and Education, Vietnam
[2]Queen's University Belfast, UK
[3]Queen Mary University of London, UK
[4]Edinburgh Napier University, UK
[5]Istanbul Technical University, Turkey
[6]Duy Tan University, Vietnam
[7]Memorial University, Canada

CORRESPONDING AUTHOR: Tan Do-Duy (e-mail: tandd@hcmute.edu.vn).

The work of Tan Do-Duy was supported by Ho Chi Minh City University of Technology and Education (HCMUTE), Vietnam, under Grant No.T2023-30. The work of B. Canberk is also supported by The Scientific and Technological Research Council of Turkey (TUBITAK) 1515 Frontier RD Laboratories Support Program for BTS Advanced AI Hub: BTS Autonomous Networks and Data Innovation Lab. Project 5239903. The work of T. Q. Duong was supported in part by the U.K. Royal Academy of Engineering (RAEng) under the RAEng Research Chair and Senior Research Fellowship scheme Grant RCSRF2021\11\4 and in part by the Canada Excellence Research Chair (CERC) Program CERC-2022-00109.

ABSTRACT In order to overcome spectrum scarcity and provide higher data rates, the sixth-generation (6G) wireless communication network is expected to perform data transmission using terahertz (THz) frequencies. However, the effective implementation of these communication systems is hampered by severe levels of signal degradation to which the THz bandwidth is subject to. Recent improvements and advancements in the fabrication process of electromagnetic (EM) metamaterials have made reconfigurable intelligent surfaces (RIS) a very promising solution to address these THz-related attenuation issues. Additionally, the adoption of non-orthogonal multiple access (NOMA) transmissions represents an effective way to improve spectrum efficiency for 6G networks. In this paper, we investigate the problem of downlink aggregated sum-rate maximisation for a multiple-input multiple-output (MIMO) system assisted by a RIS panel in performing NOMA transmission within the THz bandwidth. More specifically, we propose an optimization algorithm that jointly optimizes the transmitting power at the access point (AP) and the phase-shift coefficients for the RIS elements iteratively. Through simulation results, we demonstrate that the proposed method outperforms conventional benchmark schemes in terms of achieved aggregated throughput.

INDEX TERMS NOMA, RIS, Throughput maximisation, THz-based communications.

I. Introduction

The upcoming sixth-generation communication (6G) technologies are envisaged to provide substantial improvements in terms of communication performance, which in turn will enable the possibility of deploying a new set of disruptive services like Internet of Things based services, augmented/virtual reality, Tactile Internet, multi-sense experience and autonomous driving [1], [2]. Indeed, 6G technologies are expected to meet the stringent requirements necessary to enable the possibility to fully roll-out these services, such as massive ubiquitous connectivity, higher data rates on the order of terabits-per-second (Tbps), and ultra-reliable low latency communications (URLLCs) [3], [4]. This will be possible thanks to the innovative keystone communication paradigms and approaches that 6G will introduce, especially at the physical layer of the networks.

First of all, possibilities of performing transmissions using both sub-terahertz and terahertz (THz) bands, spanning frequencies from 100 gigahertz (GHz) to 10 THz, have been recognized by both academia and industry as essential for

the deployment of 6G-related technologies [5]. Essentially, the main interest in the adoption of these frequency bands is driven by spectrum scarcity at lower bands including sub-6 GHz, the mid-range bands, and the millimeter waves. In addition, the adoption of THz will provide some other benefits. Indeed, thanks to the advancements in major advancements in electronics and photonics, the usage of THz-based communication will enable the possibility of:

- Realizing wireless links with Tbps capacity due to the large availability of bandwidth;
- Facilitating the deployment of small-size radios that can be seamlessly embedded in diverse locations, thanks to the small dimensions of terahertz transceivers and antennas;
- Implementing novel wireless sensing techniques that extend beyond radar and localization, including climate change research and even nano-bio-sensing for revolutionary healthcare applications.

Then, it is clear how research investigations aimed to unleash the complete capabilities of THz-based communications have recently gained particular attention [6]. However, from one side the adoption of THz may open the doors to unprecedented communication capabilities, on the other hand, its adoption result is still limited. This is because THz communications signals will be subject to strong penetration losses. Although there is a vast research literature in the context of hybrid analog and digital beamforming for mitigating the channel loss problems in 5G multiple-input multiple-output (MIMO) networks [7], [8], these approaches cannot be directly applied in THz communication primarily due to the difficulty in developing highly-efficient multi-antenna transceivers for beamforming within that range of frequencies.

Interestingly, the recent progress and enhancements in the manufacturing process of electromagnetic (EM) metamaterials have led to the realization of reconfigurable intelligent surfaces (RISs), an innovative technology that permits the control of the propagation of EM waves [9]. Indeed, RIS are planar arrays of micro or nano-scale structures, usually placed on building facades, which through the usage of proper control signals exhibit various macroscopic behaviors, including reflection, refraction, and diffraction. Furthermore, RIS is a passive technology, meaning that no additional power is used to reflect the signal. For these reasons, they have been recognized as cost-effective and energy-efficient solutions for controlling and enhancing the propagation characteristics of THz-based communication scenarios [10], [11].

In addition to the usage of THz bandwidth and RIS, the concept of non-orthogonal multiple access (NOMA) represents another important communication paradigm that is gaining significant attention as a pivotal enabler in the evolution of next-generation multiple access (NGMA) techniques for 6G networks [12]–[16]. In fact, based on the concepts of superposition coding (SC) at the transmitter side, NOMA provides the possibility of multiplexing different users within the time/frequency resource block. Subsequently, receivers will be able to decode their respective message by using successive interference cancellation (SIC) [17]. Then, one can easily notice that, based on its inner principle, the adoption of NOMA holds great potential for achieving several key performance indicators in 6G networks. Indeed, the use of NOMA communication paradigm, which multiplexes multiple users within the same time-frequency resource, offers increased spectral and energy efficiency, enhanced network capacity, and improved fairness among users.

A. Related Works

Based on the previous discussion, it is clear how the adoption of the aforementioned communication paradigms and technologies will substantially contribute to meeting the 6G requirements. Numerous studies have explored the performance of each type of technology. Also, research activities and studies aimed at investigating the improvements obtained by merging those technologies have been recently presented in the literature.

The possibility of implementing a downlink THz-NOMA system has been investigated in [18]. More specifically, the authors examined a downlink communication scenario where a multiple-antenna base station employs beamforming to serve multiple clusters of users, where each beam covers a cluster of users served according to the NOMA principle. For such a communication scenario, they formulated a Beamforming-Power-Bandwidth (B-P-B) problem aimed at maximising network throughput while ensuring the quality of service (QoS) of individual users. It is noteworthy that the bandwidth allocation has been performed to capture the Long-User-Central-Window (LUCW) peculiarity of the THz band, which represents a range of THz frequencies for applications requiring high data rates over extended distances, such as wireless backhaul for 6G networks, high-speed point-to-point links, or satellite communications. Through numerical simulations, they showed how the considered THz-NOMA systems can achieve superior downlink throughput compared to equivalent Orthogonal Multiple Access (OMA) systems. Another resource allocation scheme aimed at maximising the energy efficiency (EE) of THz-NOMA systems has been proposed in [19]. In doing so, the authors first provided a channel model for the considered THz-NOMA system. Subsequently, they formulated the EE optimisation problem aimed at finding the optimal bandwidth and power allocation scheme for the users. Simulation results proved that the EE performance of THz-NOMA systems always consistently surpassed both the conventional THz Orthogonal Frequency Division Multiplexing (OFDM) system and the NOMA-without THz system.

The possibility of using RIS for assisting THz-based communication has been investigated in [20]. In this case, the authors considered a communication scenario where an

access point (AP) functions at THz frequency to facilitate indoor applications like surveillance and augmented reality (AR) services. It was assumed that RIS was placed on the ceiling of the house and used to guarantee the line-of-sight (LoS) link from the AP to indoor users. For this considered scenario, an optimisation problem aimed at finding optimal phase-shift coefficients, sub-band allocation, power allocation, and RIS deployment to maximise the system sum rate, while guaranteeing the rate requirement of each user, has been formulated and solved using a block coordinate searching (BCS) based algorithm. Authors in [21] considered the case where the usage of RIS is adopted to improve the performances of THz-based massive MIMO communication systems. More specifically, they considered the case where a multi-antenna base station (BS) intends to send data streams to a set of multi-antenna users. In this case, due to the complexity of channel estimation for the considered scenario, they proposed a ternary-tree search-based beamforming method aimed at maximising the achievable downlink rate. On the other hand, the weighted sum rate maximisation problem for a RIS-assisted multi-antenna THz-based communication system has been investigated in [22]. In this case, the authors formulated and solved the joint optimisation of BS beamforming vector and RIS reflection coefficients to maximise the weighted sum rate. Additionally, the problem of securing confidential communication in multiple RIS-assisted THz systems has been investigated in [23]. The optimisation problem involving hybrid beamforming at the BS and phase shifts at multiple RISs was defined and a corresponding optimisation algorithm was introduced to maximize the secrecy rate.

The possibility of merging the benefits of RIS-assisted communications with NOMA has also been investigated. Authors in [24] considered a RIS-aided single-input-single-output (SISO) communication network. For such a scenario, they first provided the optimal phase-shift coefficient for the RIS which maximises the strength of the received signal. Subsequently, closed-form expressions for the outage probability, ergodic rate, EE, and spectral efficiency of the considered network have been provided to show how a RIS-aided NOMA communication system permits to achieve better performances with respect to conventional NOMA. Another study aimed at highlighting the achievement of better performances in terms of reduced transmitting power to satisfy a specific user rate has been conducted in [25]. More specifically, they showed how RIS-assisted NOMA communications require less transmit power when compared to RIS-assisted OMA communications. In another paper [26], the authors analyzed the performance of a two-user uplink NOMA system augmented by a RIS. In this setup, the edge users harvest radio frequency (RF) energy from a hybrid access point (HAP) and offload their computational tasks to a server via the HAP. In addition to the analytical models, several optimization studies for RIS-assisted NOMA networks have been investigated. For instance, an energy-efficient algorithm has been proposed in [27] to find the optimal RIS phase-shift coefficients, ensuring a good tradeoff between maximising the sum rate and minimising the total power consumption. Also, in this case, the combination of RIS with NOMA achieved better performances when compared with RIS-assisted OMA. Similarly, authors in [28] illustrated how a RIS downlink transmission scheme for multiple-input single-output (MISO) systems requires less transmitting power when supports NOMA instead of only using a zero-forcing beamforming (ZFBF) scheme. In both cases, the authors provided the optimal beamforming vectors and the RIS phase shift matrix.

Studies aimed at investigating the achievable performance of a RIS-assisted THz-NOMA mobile network have been also presented. A novel THz MIMO-NOMA framework with smart reconfigurable capabilities has been investigated in [29]. In this context, the authors proposed an optimization framework aimed at maximising network energy efficiency by jointly optimizing RIS element selection and phase-shift control, as well as power allocation at the base station. The resulting highly complex optimization problem has been solved through a graph-embedded value-decomposition actor-critic based algorithm, which has been proven to outperform traditional multi-agent deep reinforcement learning algorithms. The sum-rate maximisation problem of a RIS-assisted THz-NOMA network has been investigated in [30]. More specifically, for the communication scenario under consideration, the authors proposed a new algorithm that alternatively optimises the RIS phase shift, the sub-band allocation, and power control to maximize the network's sum rate. As discussed in [31], a combined power allocation and RIS beamforming optimisation has been implemented to maximise the energy efficiency of THz-NOMA systems while ensuring users' QoS, rate fairness, and power constraints. In addition, a user pairing scheme named RTHz-NOMA has been proposed in [32]. The authors showed that the suggested user-pairing scheme considerably improves both the bit error rate and the sum rate when compared to traditional THz NOMA and OMA systems without the use of RIS. Last but not least, the authors in [33] proposed a secure transmission strategy to optimise the system secrecy rate. This was achieved by jointly optimising power allocation and phase shifts of the RIS-assisted THz-NOMA system, considering constraints such as the total transmit power, achievable rate, and RIS reflection coefficients.

B. Motivation and contributions

After a thorough review of the current literature and to the best of the author's knowledge, research activities aimed at investigating the performances of RIS-assisted THz-NOMA systems in terms of energy efficiency [29], [31], NOMA user pairing aspects [32] and system secrecy rate [33], are still limited and in their early stage. The only study closely related to our work is the one presented in [30], in which authors considered a SISO downlink communication

scenario and only one cluster of NOMA users. Motivated by the potential benefits of the joint design of NOMA and RIS resource allocation for efficient communications between the AP and distributed users in the THz band, in this paper we investigated the resource optimisation problem of downlink sum-rate maximisation through joint power allocation and phase shift selection for a massive-MIMO NOMA-THz system assisted by a RIS panel and multiple clusters of users. More specifically, the main contributions of our paper can be summarised as follows:

- We consider a downlink communication scenario involving a RIS-assisted NOMA system operating in the THz band where a RIS panel is implemented to enhance communication between a massive-MIMO AP and multiple users, which is blocked by obstacles e.g., walls. For such a scenario, we formulate a maximisation problem designed to maximise the system sum rate by jointly optimising the power control coefficients of the massive MIMO AP and the phase shifts of the RIS panel.
- We consider the case where users are clustered in different NOMA clusters according to a specific pairing scheme.
- Since the considered optimisation problem is highly non-convex, we decompose the original problem into two sub-problems. More specifically, one algorithm is for optimal AP transmit power allocation optimisation and another is for RIS phase shift optimisation, subject to the AP's power constraints and the user's QoS requirements. Then, we propose an iterative algorithm to efficiently solve the proposed optimisation problem by applying approximation and alternating optimisation (AO) methods.
- Finally, our numerical results confirm the effectiveness of the proposed resource allocation scheme in supporting the downlink RIS-assisted NOMA-THz system. Simulation results show a significant gain of the proposed scheme over the conventional schemes in terms of the sum rate of all users in the network. A complexity analysis of our proposed solution and its convergence characteristics are also provided.

The rest of the paper is organized as follows. Section II presents the considered RIS-assisted NOMA-THz communication network model, including the THz channel model, and the transmission scheme. The formulation of the sum rate maximisation problem for the considered scenario and the decomposition of the original problem into two sub-problems is illustrated in Section III. Section IV provides simulation results to demonstrate the effectiveness of the proposed method. Finally, Section V concludes the paper.

Notations: Matrices and vectors are symbolized by boldface uppercase and lowercase letters, respectively. The superscript H denotes the conjugate transpose operation of a matrix, while $\mathbb{C}^{M\times N}$ denotes the set of complex-valued matrices with of M rows and N columns.

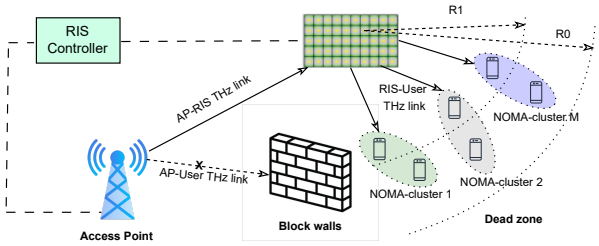


FIGURE 1: An illustration of a RIS-assisted NOMA-THz communication network.

II. System model

In this section, we will provide the considered system model and the formulation of optimisation problem for sum-rate maximisation.

A. RIS-assisted NOMA-THz network model

In Figure 1, we illustrate a downlink RIS-assisted NOMA-THz communication scenario aimed at providing service to indoor users. In this scenario, we assume that the direct links between a THz access point and users are blocked by different obstacles such as external walls. To this end, a RIS panel is deployed to assist the communication between the AP and multiple single-antenna user equipments (UEs) in the dead zone, which is the area obstructed by the external walls. We further assume that the AP is equipped with a large L-antenna array and has perfect channel state information of the entire network [27], [34]. On the other hand, we consider a RIS consisting of N discrete elements used to passively reflect the signal from the AP to a set of $\mathcal{K} = \{1, ..., K\}$ UEs. The phase-shift reflection coefficients of the RIS elements are controlled by a controller via a separate communication link [25].

The K UEs are uniformly deployed within a sphere D_0 of radius R_0 and the center at the RIS. We assume that D_0 is divided into two regions. In particular a smaller sphere D_1 with radius $R_1 < R_0$ located inside the sphere D_0 with the same center as the sphere D_0. Then, a circular region D_2 is obtained by subtracting D_1 from D_0 i.e., $D_2 = |D_0 - D_1|$ [35]. Hence, each UE located in D_1 can be randomly paired with one UE located in D_2. In this way, the original set of K UEs will be converted into a set of M NOMA clusters, each consisting of two users i.e., $K = 2M$. Note that we restrict our focus to two UEs per cluster for downlink NOMA to manage the decoding complexity involved with SIC [27]. According to these assumptions, the m-th cluster will contain 2 UEs $(m = 1, 2, ..., M)$. We will use the notation (m, i)-th to indicate the i-th UE within the m-th cluster, where $i = 1, 2$.

B. Channel model

We use a three-dimensional (3D) Cartesian coordinate system in which the 3D positions of the AP, RIS, and UEs are specified as $(0,0,0)$, (x_R, y_R, H_R), and (x_k, y_k, H_k), $k \in \mathcal{K}$, respectively. This means that the position of the AP represents the origin of the coordinate system, with the AP antenna height equal to 0, while the RIS is located at the position (x_R, y_R, H_R), representing the center of the concentrical disks D_i with $i = 0, 1, 2$, at a height H_R. Due to the high molecular absorption in THz-band communication, we assume that the LoS transmissions are more dominant than the non-line-of-sight (NLoS) transmissions [36], [37]. Therefore, the LoS channel gain between a transmitter x and a receiver y is modeled as [27]

$$h_{x,y} = \frac{c}{4\pi f d_{x,y}} e^{-k(f) d_{x,y}/2}, \tag{1}$$

where c is the speed of light; f is the operating THz frequency; $d_{x,y}$ is the distance between the transmitter and the receiver; and $k(f)$ is the frequency-dependent molecular absorption coefficient [38]. According to this model, we denote the LoS channel gain between the l-th antenna of AP and the n-th element of RIS as $h_{l,n}$, with $l = 1, 2, ..., L$, $n = 1, 2, ..., N$. Similarly, the LoS channel gain between the n-th element of RIS and the (m,i)-th UE as $h_{n,(m,i)}$.

Then, let $\mathbf{h}_l^m = [h_{l,n}^m]_{n=1}^N$ denote a $1 \times N$ vector containing the LoS channel gains $h_{l,n}^m$ between the l-th antenna of the AP and the n-th RIS element, for the sub-band allocated to the m-th cluster. We denote $\mathbf{H}_{0,m} = [\mathbf{h}_l^{mH}]_{l=1}^L \in \mathbb{C}^{N \times L}$ and $\mathbf{h}_{m,i}^H = [h_{n,(m,i)}]_{n=1}^N \in \mathbb{C}^{1 \times N}$ as the channel matrix between the AP and the RIS, and the channel vector between the RIS and the (m,i)-th UE, respectively. In this context, $[\cdot]_{first}^{last}$ represents the operation of concatenating side-by-side either vectors or single elements in the case of $\mathbf{H}_{0,m}$ and $\mathbf{h}_{m,i}^H$, respectively. Therefore, the cascaded channel matrix of the link between the AP and the (m,i)-th UE via the RIS, $\mathbf{g}_{m,i} \in \mathbb{C}^{1 \times L}$, is represented as [39]

$$\mathbf{g}_{m,i} = \mathbf{h}_{m,i}^H \mathbf{\Phi}_m \mathbf{H}_{0,m}, \tag{2}$$

where $\mathbf{\Phi}_m = diag[\phi_{1m}, \phi_{2m}, ..., \phi_{Nm}]$ represents the phase shift matrix of the RIS for the m-th cluster. Each element of this diagonal matrix represents the reflection amplitude and phase shift coefficient applied to each element of the RIS to the incident waveform [40]. These are expressed as $\phi_{nm} = \alpha_{nm} e^{j\theta_{nm}}$ with $\alpha_{nm} \in [0,1]$ and $\theta_{nm} \in [0, 2\pi]$ $(\forall n = 1, 2, ..., N,\ m = 1, ..., M)$. Based on the principle that reflecting elements do not influence the amplitude of reflected signals, as supported by [40], we can reasonably assume $\alpha_{nm} = 1$. In addition, we define $\mathbf{\Phi} = [\mathbf{\Phi}_m]_{m=1}^M$ as the phase shifts of the RIS with respect to all clusters obtained by concatenating the diagonal matrices $\mathbf{\Phi}_m$ $(m = 1, ..., M)$ side by side, which results in a matrix of N rows and $L \times M$ columns.

C. Transmission scheme

In terms of the transmission model, we assume that the total bandwidth B is divided into M equal sub-bands of size $\frac{B}{M}$, each of them allocated to each NOMA cluster, i.e. one frequency slot per NOMA cluster. For the UEs in NOMA clusters, we assume that the perfect SIC is exploited for NOMA users [41], [42], meaning that the user with a higher channel coefficient can successfully remove the signal of the user with a lower channel coefficient. Without loss of generality, assume that the channel gain of the two UEs of the m-th NOMA cluster is sorted $|\mathbf{g}_{m,1}\mathbf{f}_{m,1}| \geq |\mathbf{g}_{m,2}\mathbf{f}_{m,2}|$ where $\mathbf{g}_{m,1}$ and $\mathbf{g}_{m,2}$ are the cascaded channel matrices of the links between the AP and the UEs of the m-th NOMA cluster given by (2), while $\mathbf{f}_{m,1} = \frac{\mathbf{g}_{m,1}^H}{\|\mathbf{g}_{m,1}\|}$ and $\mathbf{f}_{m,2} = \frac{\mathbf{g}_{m,2}^H}{\|\mathbf{g}_{m,2}\|}$ are the transmit beamforming vectors of the AP for each UE in the m-th cluster. The AP is assumed to adopt the optimal beamforming vector to maximize the signal-to-noise ratio. Let $\mathbf{p}_m = \{\alpha_1^m, \alpha_2^m\}$, with $\alpha_1^m + \alpha_2^m = 1$, denote the fraction of the power of the AP allocated for the first and second users of the m-th NOMA cluster, respectively. We indicate the entire matrix of power coefficients with $\mathbf{p} = [\mathbf{p}_m]_{m=1}^M$, obtained.

According to the previous notation, the achievable rate for the user experiencing strong channel gain within the m-th NOMA cluster is:

$$R_{m,1}(\mathbf{p}_m, \mathbf{\Phi}_m) = \frac{B}{M} \log_2 \left(1 + \frac{\alpha_1^m P_0 |\mathbf{g}_{m,1}\mathbf{f}_{m,1}|^2}{\sigma^2}\right). \tag{3}$$

The achievable rate for the user experiencing weak channel gain within the m-th NOMA cluster is:

$$R_{m,2}(\mathbf{p}_m, \mathbf{\Phi}_m) = \frac{B}{M} \log_2 \left(1 + \frac{\alpha_2^m P_0 |\mathbf{g}_{m,2}\mathbf{f}_{m,2}|^2}{\alpha_1^m P_0 |\mathbf{g}_{m,2}\mathbf{f}_{m,1}|^2 + \sigma^2}\right), \tag{4}$$

where P_0 is the transmit power budget of the AP and σ^2 is the white power spectral density at the users. Hence, the total achievable rate of the network is given by

$$R_{tot}^{noma}(\mathbf{p}, \mathbf{\Phi}) = \sum_{m=1}^{M} \sum_{i=1}^{2} R_{m,i}(\mathbf{p}_m, \mathbf{\Phi}_m). \tag{5}$$

We compare the performance of the proposed NOMA system with that of the traditional Frequency Division Multiple Access (FDMA) system where each UE is assigned to a specific bandwidth. Let $\mathbf{g}_k \in \mathbb{C}^{1 \times L}$ denote the cascaded channel matrix of the link between the AP and the k-th UE via the RIS given by

$$\mathbf{g}_k = \mathbf{h}_k^H \mathbf{\Phi}_k \mathbf{H}_{0,k}, \tag{6}$$

where $\mathbf{h}_l^k = [h_{l,n}^k]_{n=1}^N$ represents a $1 \times N$ vector that encapsulates the LoS channel gains $h_{l,n}^k$ between the l-th antenna of the AP and the n-th RIS element for the sub-band designated to the k-th UE; $\mathbf{H}_{0,k} = [\mathbf{h}_l^{kH}]_{l=1}^L$ and $\mathbf{h}_k^H = [h_{n,k}]_{n=1}^N$ denote the channel matrix between the AP and the RIS, and the channel vector between the RIS and the k-th UE,

respectively; $\mathbf{\Phi}_k = diag[\phi_{1k}, \phi_{2k}, ..., \phi_{Nk}]$ is the phase shift matrix of the RIS for the k-th UE with $\phi_{nk} = \alpha_{nk} e^{j\theta_{nk}}$ ($\alpha_{nk} \in [0,1]$ and $\theta_{nk} \in [0, 2\pi]$); and $\mathbf{\Phi}^{'} = [\mathbf{\Phi}_k]_{k=1}^{K}$ is the phase shifts of the RIS for all UEs, constructed by placing the diagonal matrices $\mathbf{\Phi}_k$ $(k = 1, ..., K)$ adjacent to each other. Consequently, the total achievable rate of the FDMA-based system is given by

$$R_{tot}^{fdma}(\mathbf{p}, \mathbf{\Phi'}) = \sum_{k=1}^{K} \frac{B}{K} \log_2 \left(1 + \frac{P_0 |\mathbf{g}_k \mathbf{f}_k|^2}{\sigma^2}\right), \quad (7)$$

where $\mathbf{f}_k = \frac{\mathbf{g}_k^H}{\|\mathbf{g}_k\|}$ denotes the transmit beamforming vectors of the AP for the k-th UE.

D. Problem formulation

In this paper, we aim to maximise the total achievable rate of all UEs in the network by exploiting the downlink NOMA-THz network and by jointly optimising the power allocation coefficients at the AP ($\mathbf{p}$) and the phase shifts of the RIS panel ($\mathbf{\Phi}$), subject to the QoS requirements at the UEs. To this end, the optimisation problem is explicitly formulated as follows

$$\mathsf{P1}: \max_{\mathbf{p},\ \mathbf{\Phi}} \ R_{tot}^{noma}(\mathbf{p}, \mathbf{\Phi}) \quad (8a)$$

$$\text{s.t. } \alpha_1^m + \alpha_2^m = 1, \ \forall m = 1 \cdots M \quad (8b)$$

$$R_{m,1}(\mathbf{p}_m, \mathbf{\Phi}_m) \geq \hat{r}_{m,1}, \ \forall m = 1 \cdots M; \quad (8c)$$

$$R_{m,2}(\mathbf{p}_m, \mathbf{\Phi}_m) \geq \hat{r}_{m,2}, \forall m = 1 \cdots M; \quad (8d)$$

$$0 \leq \theta_{nm} \leq 2\pi, \forall \ n = 1 \cdots N, \forall m = 1 \cdots M, \quad (8e)$$

where the constraint (8b) specifies the range of power allocation coefficients. The constraints (8c)-(8d) formulate the QoS requirements at the corresponding UEs, with $\hat{r}_{m,1}$ and $\hat{r}_{m,2}$ being the minimum achievable rate required at the UEs. The constraint (8e) shows the range of phase shifts for the n-th reflecting elements of the RIS panel.

III. Joint resource allocation optimisation problem

The joint optimisation problem formulated in problem P1 results in being highly non-convex to solve. In this section, we illustrate our proposed approach consisting of decomposing the original problem into two iterative sub-problems. One problem will aim at determining the optimal power allocation for a fixed set of RIS coefficients. Subsequently, another optimisation problem is formulated to find the optimal set of RIS coefficients for a fixed value of power allocation coefficients. The solutions proposed for each sub-problem are executed iteratively until a convergence condition is met.

A. Power Allocation Optimisation

For a fixed value of RIS coefficients in problem P1, we obtain the following power allocation optimisation problem:

$$\mathsf{P2}: \max_{\mathbf{p}} \ R_{tot}^{noma}(\mathbf{p}) \quad (9a)$$

$$\text{s.t. } \alpha_1^m + \alpha_2^m = 1, \ \forall m = 1 \cdots M; \quad (9b)$$

$$R_{m,1}(\mathbf{p}_m) \geq \hat{r}_{m,1}, \ \forall m = 1 \cdots M; \quad (9c)$$

$$R_{m,2}(\mathbf{p}_m) \geq \hat{r}_{m,2}, \ \forall m = 1 \cdots M. \quad (9d)$$

This is a maximisation problem having a concave objective function and convex constraints, which becomes challenging to solve due to the logarithm functions in (9a) and (9c)-(9d). To overcome this complexity in solving problem P2, we adopt an efficient approximation approach already adopted in [43], [44], which consists of exploiting the logarithmic inequalities (see Appendix A for detailed proofs). By using efficient logarithmic inequality approximations, the logarithm functions in (9a) and constraints (9c)-(9d) can be replaced by a lower bound approximation which is more tractable or easier to solve using convex optimisation solvers such as CVX [45].

Hence, at the ψ-th iteration of the algorithm, problem P2 can be rewritten as

$$\mathsf{P3}: \max_{\mathbf{p}^{(\psi)}} \ \hat{R}_{tot}^{noma}(\mathbf{p}^{(\psi)}) \quad (10a)$$

$$\text{s.t. } \alpha_1^m + \alpha_2^m = 1, \ \forall m = 1 \cdots M; \quad (10b)$$

$$\hat{R}_{m,1}^{(\psi)}\left(\mathbf{p}_m^{(\psi)}\right) \geq \hat{r}_{m,1}, \ \forall m = 1 \cdots M; \quad (10c)$$

$$\hat{R}_{m,2}^{(\psi)}\left(\mathbf{p}_m^{(\psi)}\right) \geq \hat{r}_{m,2}, \ \forall m = 1 \cdots M. \quad (10d)$$

where $\hat{R}_{m,1}^{(\psi)}\left(\mathbf{p}_m^{(\psi)}\right)$ and $\hat{R}_{m,2}^{(\psi)}\left(\mathbf{p}_m^{(\psi)}\right)$ are defined as (25) and (26), respectively, and

$$\hat{R}_{tot}^{noma}(\mathbf{p}^{(\psi)}) = \sum_{m=1}^{M} \sum_{i=1}^{2} \hat{R}_{m,i}^{(\psi)}\left(\mathbf{p}_m^{(\psi)}\right). \quad (11)$$

By using the optimisation tool CVX, we can efficiently solve problem P3 with a computational complexity of $\mathcal{O}((3M)^{3.5})$ [46].

In Algorithm 1, we introduce an AO-based algorithm to solve the power allocation problem P3.

Algorithm 1 : Proposed algorithm to solve problem P3.

Input:
Set initial iteration number $\psi = 0$
Set initial phase shifts $\mathbf{\Phi}$ and initial power control coefficients $\mathbf{p}^{(0)}$
Set the maximum number of iterations $I_{max} = 20$
Set the tolerance $\varepsilon = 10^{-3}$
Repeat
Solve problem P3
Obtain the current optimal coefficients $(\mathbf{p}^{(\psi+1)})$
Set $\psi = \psi + 1$
Until Convergence or the iteration reaches I_{max}
Output: Optimal power control coefficients $(\mathbf{p}^*)$

B. Phase Shift Optimisation

Given any set of power allocation coefficients $\mathbf{p}$, problem P1 can be restructured as

$$\mathsf{P4}: \max_{\mathbf{\Phi}} \; R_{tot}^{noma}(\mathbf{\Phi}) \tag{12a}$$
$$\text{s.t. } R_{m,1}\left(\mathbf{\Phi}_m\right) \geq \hat{r}_{m,1}, \; \forall m = 1 \cdots M; \tag{12b}$$
$$R_{m,2}\left(\mathbf{\Phi}_m\right) \geq \hat{r}_{m,2}, \; \forall m = 1 \cdots M; \tag{12c}$$
$$0 \leq \theta_{nm} \leq 2\pi, \; \forall n = 1 \cdots N. \tag{12d}$$

By considering the following vector of variables $\boldsymbol{\eta_m} = [\eta_m^1, ..., \eta_m^N]^H$ with $\eta_m^n = e^{j\theta_{nm}}$, the constraint (12d) can be rewritten as unit-modulus constraint i.e., $|\eta_m^n|^2 = 1$ [47]. In addition, by another variable $\Gamma_{m,i} = diag(\mathbf{h}_{m,i}^H)\mathbf{H}_{0,m}\mathbf{f}_{m,i}$ $(\forall i = 1, 2)$, one case easily notices that $\mathbf{g}_{m,i}\mathbf{f}_{m,i} = \boldsymbol{\eta_m}^H \Gamma_{m,i}$. Problem P4 is partitioned into M separate subproblems, with each subproblem associated with a distinct cluster. Then, at the j-th iteration of the algorithm, problem P4 can be rewritten as:

$$\mathsf{P5.m}: \max_{\boldsymbol{\eta_m}^{(j)}} \; \hat{R}_{tot}^{noma}(\boldsymbol{\eta_m}^{(j)}) \tag{13a}$$
$$\text{s.t. } \boldsymbol{\eta_m}^H \Gamma_{m,1}\Gamma_{m,1}^H \boldsymbol{\eta_m} \geq \left(2^{M\hat{r}_{m,1}} - 1\right)/a_1^m, \tag{13b}$$
$$\boldsymbol{\eta_m}^H \Gamma_{m,2}\Gamma_{m,2}^H \boldsymbol{\eta_m} \geq \left(2^{M\hat{r}_{m,2}} - 1\right)\left(\alpha \boldsymbol{\eta_m}^H \Gamma_{m,1}\Gamma_{m,1}^H \boldsymbol{\eta_m} + \frac{1}{a_2^m}\right), \tag{13c}$$
$$|\eta_m^n|^2 = 1, \forall n = 1 \cdots N, \tag{13d}$$

where P#.m denotes a per m-th cluster optimization problem, $a_1^m = \alpha_1^m P_0/\sigma^2$, $a_2^m = \alpha_2^m P_0/\sigma^2$, and $\alpha = \alpha_1^m/\alpha_2^m$, and

$$\hat{R}_{tot}^{noma}(\boldsymbol{\eta_m}^{(j)}) = \sum_{m=1}^{M}\sum_{i=1}^{2} \hat{R}_{m,i}\left(\boldsymbol{\eta_m}^{(j)}\right). \tag{14}$$

with

$$\hat{R}_{m,1} = \frac{B}{M}\log_2\left(1 + a_1^m \boldsymbol{\eta_m}^H \Gamma_{m,1}\Gamma_{m,1}^H \boldsymbol{\eta_m}\right),$$

$$\hat{R}_{m,2} = \frac{B}{M}\log_2\left(1 + \frac{a_2^m \boldsymbol{\eta_m}^H \Gamma_{m,2}\Gamma_{m,2}^H \boldsymbol{\eta_m}}{a_1^m \boldsymbol{\eta_m}^H \Gamma_{m,1}\Gamma_{m,1}^H \boldsymbol{\eta_m} + 1}\right).$$

However, this new equivalent problem is a non-convex quadratically constrained quadratic programming (QCQP) problem. To find a solution, we employ the following transformation. Let us define $\mathbf{U}_{m,i} = \Gamma_{m,i}\Gamma_{m,i}^H$ and $\boldsymbol{\eta_m}^H \mathbf{U}_{m,i} \boldsymbol{\eta_m} = tr\left(\mathbf{U}_{m,i}\mathbf{Z}_m\right)$ $(\forall i = 1, 2)$ where $\mathbf{Z}_m = \boldsymbol{\eta_m}\boldsymbol{\eta_m}^H$ must satisfy $\mathbf{Z}_m \succeq \mathbf{0}$ and $rank(\mathbf{Z}_m) = 1$. At this point, the rank-one constraint can be relaxed because of the non-convex nature of the unit-modulus constraint (13d) [47], allowing to write problem P5.m in the following new equivalent form:

$$\mathsf{P6.m}: \max_{\boldsymbol{\eta_m}^{(j)}} \; \hat{R}_{tot}^{noma}(\boldsymbol{\eta_m}^{(j)}) \tag{15a}$$
$$\text{s.t. } tr\left(\mathbf{U}_{m,1}\mathbf{Z}_m\right) \geq \left(2^{K\hat{r}_{m,1}/2} - 1\right)/a_1^m, \tag{15b}$$
$$tr\left(\mathbf{U}_{m,2}\mathbf{Z}_m\right) \geq \left(2^{K\hat{r}_{m,2}/2} - 1\right)\left(\alpha tr\left(\mathbf{U}_{m,1}\mathbf{Z}_m\right) + \frac{1}{a_2^m}\right), \tag{15c}$$
$$\mathbf{Z}_m(n,n) = 1, \forall n = 1 \cdots N, \tag{15d}$$
$$\mathbf{Z}_m \succeq \mathbf{0}, \tag{15e}$$

where

$$\hat{R}_{tot}^{noma}(\boldsymbol{\eta_m}^{(j)}) = \sum_{m=1}^{M}\sum_{i=1}^{2} \hat{R}_{m,i}\left(\boldsymbol{\eta_m}^{(j)}\right). \tag{16}$$

with

$$\hat{R}_{m,1} = \frac{B}{M}\log_2\left(1 + a_1^m tr\left(\mathbf{U}_{m,1}\mathbf{Z}_m\right)\right),$$

$$\hat{R}_{m,2} = \frac{B}{M}\log_2\left(1 + \frac{a_2^m tr\left(\mathbf{U}_{m,2}\mathbf{Z}_m\right)}{a_1^m tr\left(\mathbf{U}_{m,1}\mathbf{Z}_m\right) + 1}\right),$$

Follow the approximation approach in (20), at the jth iteration, we can replace $\hat{R}_{m,2}$ by the following lower bound as

$$\hat{R}_{m,2} \geq \tilde{R}_{m,2}^{(j)}, \tag{17}$$

where

$$\tilde{R}_{m,2}^{(j)} = \frac{B}{M}\Big[\log_2\left(1 + \frac{1}{\bar{u}_2\bar{v}_2}\right) + \frac{2}{\bar{u}_2\bar{v}_2 + 1} - \frac{u_2}{\bar{u}_2(\bar{u}_2\bar{v}_2 + 1)} - \frac{v_2}{\bar{v}_2(\bar{u}_2\bar{v}_2 + 1)}\Big], \tag{18}$$

with

$$u_2 = \frac{1}{a_2^m tr\left(\mathbf{U}_{m,2}\mathbf{Z}_m\right)}, v_2 = a_1^m tr\left(\mathbf{U}_{m,1}\mathbf{Z}_m\right) + 1,$$

$$\bar{u}_2 = u_2^{(j)} = \frac{1}{a_2^m tr\left(\mathbf{U}_{m,2}^{(j)}\mathbf{Z}_m\right)},$$

$$\bar{v}_2 = v_2^{(j)} = a_1^m tr\left(\mathbf{U}_{m,1}^{(j)}\mathbf{Z}_m\right) + 1.$$

As seen, problem P6 is now a convex semi-definite program (SDP) [47], allowing it to be efficiently solved by convex optimisation solvers such as CVX.

In Algorithm 2, we propose a solution procedure based on the AO algorithm to solve problem P6. In particular, the procedure is designed to identify the optimal phase shifts according to each UE cluster and terminates once the optimal phase shifts of all clusters have converged. Based on the complexity analysis provided in Section III-A, the computational complexity of Algorithm 2 is $\mathcal{O}((2M(N+1))^{2.5})$.

Algorithm 2 : Proposed AO-based algorithm to solve problem P6.

Input:
Set initial iteration number $j = 0$
Set initial power control coefficients $\mathbf{p}$ and initial phase shifts $\mathbf{\Phi}^{(0)}$
Set the maximum number of iterations $I_{max} = 20$
Set the tolerance $\varepsilon = 10^{-3}$
Repeat
for $m = 1$ to M
Solve problem P6.m
Obtain the current optimal phase shifts $(\mathbf{\Phi}^{(j+1)})$
end for
Set $j = j + 1$.
Until Convergence or the iteration reaches I_{max}.
Output: Optimal RIS phase shifts $(\mathbf{\Phi}^*)$

C. Iterative Optimisation Algorithm

In Sections III-A and III-B , the entire optimisation problem P1 has been decomposed into two convex problems P3 and P6. Subsequently, we propose an iterative algorithm to solve problem P1 as in Algorithm 3. In particular, Algorithm 3 sets the initial iteration number $\zeta = 0$, and takes the initial phase shifts coefficients $\mathbf{\Phi}^{(0)}$ and the initial power control coefficients $\mathbf{p}^{(0)}$. At each iteration, these are updated using Algorithm 1 and Algorithm 2, respectively. In other words, Algorithm 3 summarizes the optimisation flow where the solution from each iteration serves as the starting point of the next iteration. Algorithm 3 is terminated either when the maximum number of iterations I_{max} is reached, or the difference between the solutions within two consecutive iterations is less than ε, which are both set during the initialization phase of the algorithm. Furthermore, one can note how Algorithm 2 is significantly more complex than Algorithm 1. More specifically, the computational complexity of Algorithm 3 is $I \times \mathcal{O}((2M(N+1))^{2.5})$, where I is the number of required iterations.

Algorithm 3 Iterative algorithm to jointly solve problem P1

Input:
Set initial iteration number $\zeta = 0$
Set initial power control coefficients $\mathbf{p}^{(0)}$ and initial phase shifts $\mathbf{\Phi}^{(0)}$
Set the maximum number of iterations $I_{max} = 20$
Set the tolerance $\varepsilon = 10^{-3}$
repeat
Solve problem P3 using Algorithm 1 to obtain the current solution $(\mathbf{p}^{(\zeta+1)})$
Solve problem P6 using Algorithm 2 to obtain the current solution $(\mathbf{\Phi}^{(\zeta+1)})$
Set $\zeta = \zeta + 1$
until Convergence or the iteration reaches I_{max}
Output: Optimal solution $(\mathbf{p}^*, \mathbf{\Phi}^*)$

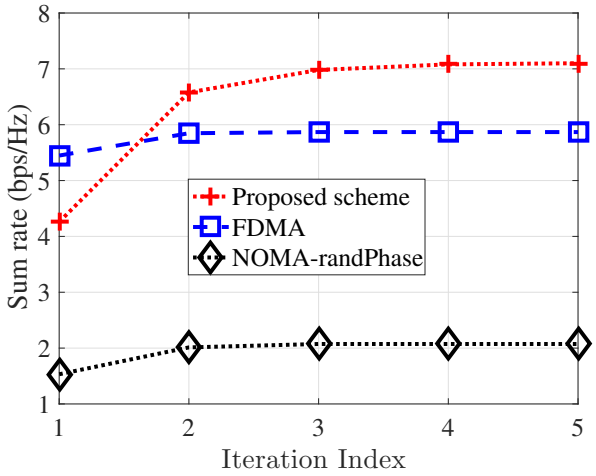


FIGURE 2: The convergence of Algorithm 3 for solving the proposed scheme, FDMA, and NOMA-randPhase scheme, at $K = 4$, $N = 128$, $P_0 = 40dbm$.

IV. Performance Evaluation

This section provides the results from the numerical simulations carried out to assess the performance of the proposed optimisation algorithm. In doing so, it has been supposed that the AP and RIS are located at coordinates (0, 0, 2) and (5, 5, 2) meters, respectively. This means that the AP antenna and RIS panels have the same height of 2 meters. The UEs are uniformly distributed in a circle within a circle 10 meters radius centered at (10, 10, 0) meters. On the other hand, the radius of the sphere D_0 and D_1 are $R_0 = 5\ m$ and $R_1 = 2.5\ m$, respectively. The AP performs transmission within a $B = 10$ GHz bandwidth with central carrier frequency set at $f_c = 300$ GHz and white noise power spectral density equal to $\sigma^2 = -174$ dBm/Hz [19]. The minimum achievable rate required at the UEshas been set to $r_{min} = 0.1$ bps/Hz [20]. As regards the optimisation algorithm, the tolerance threshold and the iterations for convergence have been set to $\epsilon = 10^{-3}$ and $I_{max} = 10$, respectively. An extensive simulation campaign has been carried out by varying the power budget at the AP and the number of passive reflecting elements of the RIS, while the AP has been supposed to be equipped with 8 antennas. All results are averaged over multiple simulation rounds according to different random realizations of UE locations. The performances of the proposed optimisation solution have been compared with the following benchmark schemes:

- FDMA scheme: In this case only the phase-shift coefficients of the RIS are optimized, while the power allocation scheme is not necessary, i.e. only Algorithm 2 is used and the maximum transmitting power is assigned to each user.

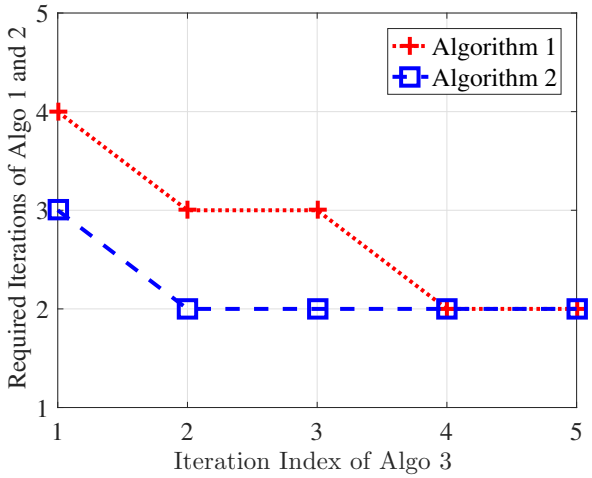


FIGURE 3: The convergence of Algorithm 1 and Algorithm 2 with respect to the convergence of Algorithm 3 for solving the proposed scheme at $K = 4$, $N = 128$, $P_0 = 40dbm$.

- NOMA-randPhase: The same power allocation methodology as the proposed in Algorithm 1 while the phase-shift coefficients of RIS elements are randomly allocated at the beginning.

A. Convergence of the proposed algorithms

Figure 2 denotes the convergence characteristic of Algorithm 3 for solving problem P1 at $K = 4$, $N = 128$, $P_0 = 40dbm$, as well as to find the optimal resource allocation for the other benchmark schemes. From this figure, it is possible to observe that our proposed optimisation scheme demands only a few iterations to identify the optimal resource allocation for each transmission scheme considered. More specifically, for the FDMA and NOMA-randPhase schemes, the algorithm reaches the optimal solution after 3 iterations. On the other hand, it requires more iterations in the case of joint optimisation. This is justified by the higher number of variables that Algorithm 3 has to consider to perform a joint optimisation. However, the most relevant result is the higher value of the downlink sum rate achieved when the power coefficients and phase shifts are jointly optimised. Furthermore, in Figure 3, we also show the convergence in terms of the number of iterations that Algorithms 1 and 2 need to solve problems P3 and P6, respectively, according to each iteration of Algorithm 3. As denoted in (3) and (4), the power levels allocated to the two users in each NOMA cluster are interdependent. Therefore, adjusting the power allocation for one user affects the interference levels experienced by the other. Hence, this coupling adds complexity to the optimization problem P3 when compared with problem P6.

B. System sum rate versus the number of RIS elements

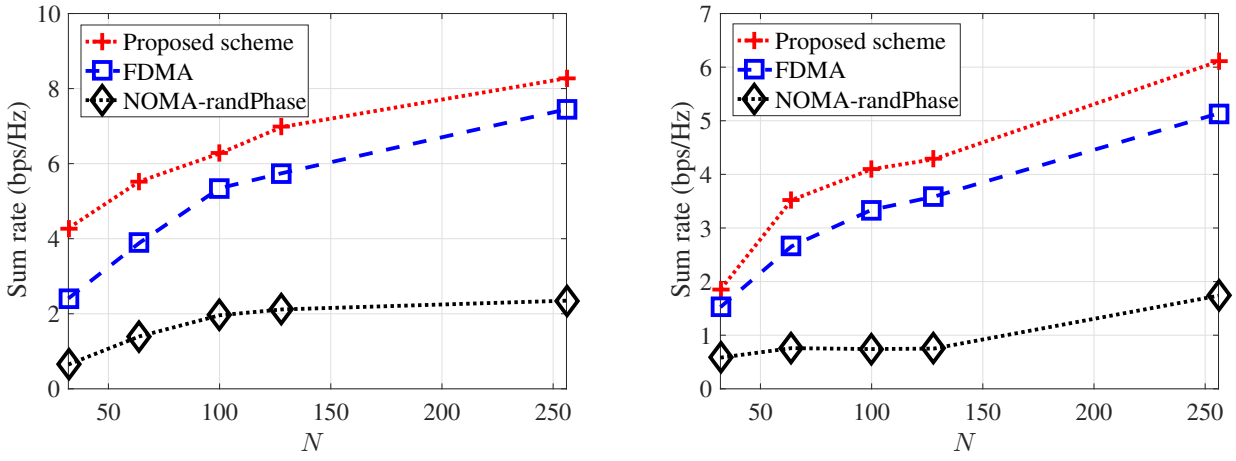
Figure 4 illustrates the average sum rate as a function of different numbers of RIS elements (N) and for a fixed value of P_0. This figure demonstrates that for all the considered schemes, the sum rate achieved increases as the number of RIS elements increases. This represents clear evidence of the benefits of using RIS to improve the characteristics of wireless communication channels. Indeed, the higher the number of reflecting elements, the higher the power that the UEs can receive from the AP, thus leading to higher levels of achieved rate. In addition, these figures also illustrate how the considered RIS-assisted NOMA-THz network with the proposed scheme always outperforms the other communication schemes. In contrast to the FDMA with optimal phase shifts, the adoption of NOMA allows users to use more bandwidth, increasing the achievable rate. On the other hand, compared to the NOMA-randPhase scheme, the RIS-assisted NOMA-THz network utilizing the proposed scheme consistently outperforms the network with random phase shifts, highlighting then the importance of optimizing the phase-shift coefficients. Last but not least, one can also easily notice that by increasing the number of clusters (K) the maximum achievable rate decreases. This is justified by the fact that the portion of bandwidth decreases. However, the proposed scheme still succeeds in achieving the maximum. Specifically, compared with FDMA and NOMA-randPhase, the proposed scheme can obtain a sum rate gain up to 11% and 250% for $K = 4$, $N = 256$ (Figure 4(a)) and 19% and 245% for $K = 8$, $N = 256$ (Figure 4(b)), respectively.

C. System sum rate versus transmit power budget

Figure 5 depicts the impact of the transmit power budget (P_0) on the average system sum rate. As expected, the sum rate improves with an increase in the power budget. Additionally, the proposed scheme outperforms both the FDMA and NOMA-randPhase schemes. This result again confirms the superiority of the NOMA-based approach on spectrum efficiency when compared with the FDMA scheme. Indeed, in both cases, the phase-shift coefficients are optimised, but the adoption of NOMA allows to allocate more bandwidth to the users, even when the number of UEs increases. Specifically, as illustrated in Figure 5, when the power budget ranges from 30 to 40 dBm, the average sum rate gain obtained with our proposed scheme can reach up to 210% and 68% with $K = 4$ and $K = 8$, respectively.

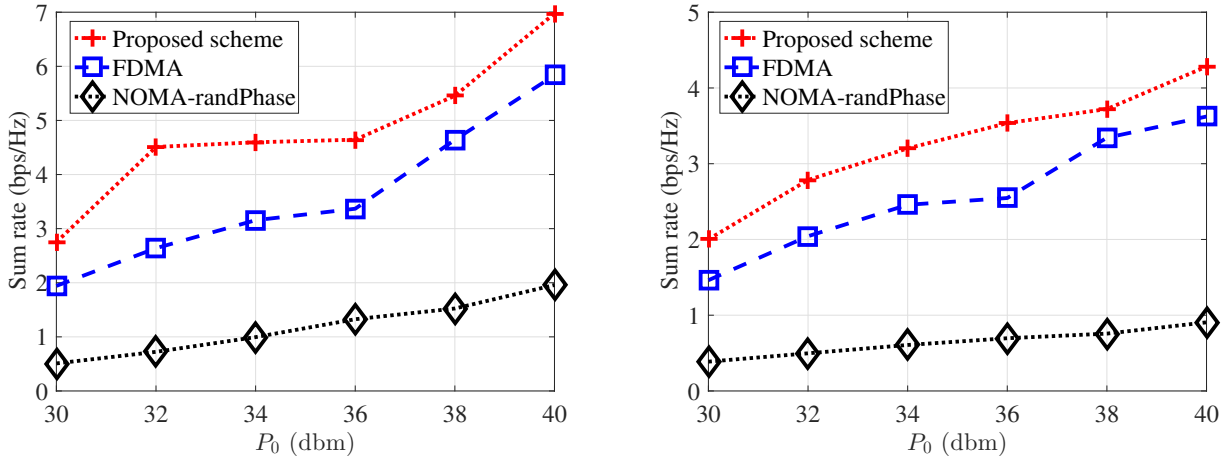
V. Conclusion

In line with the current research direction in the areas of THz-based communication for increased network capacity and data rate, adoption of NOMA for improved spectrum efficiency, and RIS for improving the propagation characteristics of the wireless channel, in this paper, we considered a MIMO-based and RIS-assisted NOMA-THz communication scenario. More specifically we formulated an optimisation problem aimed at maximizing the aggregated downlink



(a) K=4 (b) K=8

FIGURE 4: Average sum rate of all UEs versus different number of RIS elements (N), at $P_0 = 40$ dbm.



(a) K=4 (b) K=8

FIGURE 5: Average sum rate of all UEs versus different values of transmit power budget (P_0), at $N = 128$.

throughput by jointly optimizing the transmitting power at the AP and phase-shift coefficients for the RIS panel. Due to the complexity of this optimisation problem, we decomposed the entire problem into two separate and more easy sub-problems. In particular, one problem for power allocation and another for phase-shift coefficient optimisation, each of them solved iteratively with the usage of CVX solver. Through an extensive simulation campaign, we proved the superiority of our optimisation scheme when compared with other conventional schemes, as well as how the usage of NOMA and RIS panels can substantially contribute to increasing the network capacity.

Appendix A
Approximation approaches and logarithmic inequalities used to solve optimisation problems P2

To solve problem P2, we adopt the approximation approaches of [43], [44], which exploits the following logarithmic inequalities:

$$f(w) = \log_2(1 + \frac{1}{w}) \geq \hat{f}(w), \tag{19}$$

$$f(u, v) = \log_2(1 + \frac{1}{uv}) \geq \hat{f}(u, v), \tag{20}$$

in which:

$$\hat{f}(w) = \log_2\left(1 + \frac{1}{\bar{w}}\right) + \frac{1}{\bar{w}+1} - \frac{w}{\bar{w}\,(\bar{w}+1)}, \tag{21}$$

$$\begin{aligned}\hat{f}(u, v) = &\log_2\left(1 + \frac{1}{\bar{u}\bar{v}}\right) + \frac{2}{\bar{u}\bar{v}+1} \\ &- \frac{u}{\bar{u}(\bar{u}\bar{v}+1)} - \frac{v}{\bar{v}(\bar{u}\bar{v}+1)}.\end{aligned} \tag{22}$$

The inequalities in (19)-(22) are valid $\forall w > 0, \bar{w} > 0, \forall u > 0, \bar{u} > 0, v > 0, \bar{v} > 0$. At this point, by defining the following quantities at the ψ-th iteration of the algorithm:

$$w_1 = \frac{\sigma^2}{\alpha_1^m P_0 |\mathbf{G}_{m,1}\mathbf{f}_{m,1}|^2},$$

$$u_1 = \frac{1}{\alpha_2^m P_0 |\mathbf{G}_{m,2}\mathbf{f}_{m,2}|^2}, v_1 = \alpha_1^m P_0 |\mathbf{G}_{m,2}\mathbf{f}_{m,2}|^2 + \sigma^2,$$

$$\bar{w}_1 = w_1^{(\psi)} = \frac{\sigma^2}{\alpha_1^{m^{(\psi)}} P_0 |\mathbf{G}_{m,1}\mathbf{f}_{m,1}|^2},$$

$$\bar{u}_1 = u_1^{(\psi)} = \frac{1}{\alpha_2^{m^{(\psi)}} P_0 |\mathbf{G}_{m,2}\mathbf{f}_{m,2}|^2},$$

$$\bar{v}_1 = v_1^{(\psi)} = \alpha_1^{m^{(\psi)}} P_0 |\mathbf{G}_{m,2}\mathbf{f}_{m,2}|^2 + \sigma^2,$$

we derive the following approximation for the achievable rate in (3) (4), respectively, as follows:

$$R_{m,1}(\mathbf{p}_m) \geq \hat{R}_{m,1}^{(\psi)}(\mathbf{p}_m), \tag{23}$$

$$R_{m,2}(\mathbf{p}_m) \geq \hat{R}_{m,2}^{(\psi)}(\mathbf{p}_m), \tag{24}$$

where

$$\hat{R}_{m,1}^{(\psi)}(\mathbf{p}_m) = \frac{B}{M}\left[\log_2\left(1 + \frac{1}{\bar{w}_1}\right) + \frac{1}{\bar{w}_1 + 1} - \frac{w_1}{\bar{w}_1(\bar{w}_1 + 1)}\right], \tag{25}$$

$$\hat{R}_{m,2}^{(\psi)}(\mathbf{p}_m) = \frac{B}{M}\left[\log_2\left(1 + \frac{1}{\bar{u}_1\bar{v}_1}\right) + \frac{2}{\bar{u}_1\bar{v}_1 + 1} - \frac{u_1}{\bar{u}_1(\bar{u}_1\bar{v}_1 + 1)} - \frac{v_1}{\bar{v}_1(\bar{u}_1\bar{v}_1 + 1)}\right]. \tag{26}$$

Acknowledgement

The work of Tan Do-Duy was supported by Ho Chi Minh City University of Technology and Education (HCMUTE), Vietnam, under Grant No.T2023-30. The work of B. Canberk is also supported by The Scientific and Technological Research Council of Turkey (TUBITAK) 1515 Frontier RD Laboratories Support Program for BTS Advanced AI Hub: BTS Autonomous Networks and Data Innovation Lab. Project 5239903. The work of T. Q. Duong was supported in part by the U.K. Royal Academy of Engineering (RAEng) under the RAEng Research Chair and Senior Research Fellowship scheme Grant RCSRF2021\11\4 and in part by the Canada Excellence Research Chair (CERC) Program CERC-2022-00109.

REFERENCES

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

[2] W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, "The road towards 6g: A comprehensive survey," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 334–366, 2021.

[3] B. Zong, C. Fan, X. Wang, X. Duan, B. Wang, and J. Wang, "6G technologies: Key drivers, core requirements, system architectures, and enabling technologies," *IEEE Veh. Technol. Mag.*, vol. 14, no. 3, pp. 18–27, 2019.

[4] D. Van Huynh, S. R. Khosravirad, A. Masaracchia, O. A. Dobre, and T. Q. Duong, "Edge intelligence-based ultra-reliable and low-latency communications for digital twin-enabled metaverse," *IEEE Wireless Communications Letters*, vol. 11, no. 8, pp. 1733–1737, 2022.

[5] N. U. Saqib, M. S. Haroon, H. Y. Lee, K. Park, H.-G. Song, and S.-W. Jeon, "Thz communications: A key enabler for future cellular networks," *IEEE Access*, vol. 11, pp. 117 474–117 493, 2023.

[6] W. Jiang, Q. Zhou, J. He, M. A. Habibi, S. Melnyk, M. El-Absi, B. Han, M. D. Renzo, H. D. Schotten, F.-L. Luo, T. S. El-Bawab, M. Juntti, M. Debbah, and V. C. M. Leung, "Terahertz communications and sensing for 6G and beyond: A comprehensive review," *IEEE Commun. Surveys Tuts.*, pp. 1–1, 2024.

[7] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Select. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[8] I. Ahmed, H. Khammari, A. Shahid, A. Musa, K. S. Kim, E. D. Poorter, and I. Moerman, "A survey on hybrid beamforming techniques in 5G: Architecture and system model perspectives," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 3060–3097, 2018.

[9] G. C. Alexandropoulos, G. Lerosey, M. Debbah, and M. Fink, "Reconfigurable intelligent surfaces and metamaterials: The potential of wave propagation control for 6G wireless communications," *IEEE ComSoc TCCN Newslett.*, vol. 6, no. 1, pp. 25–37, Jun. 2020.

[10] E. Calvanese Strinati, G. C. Alexandropoulos, H. Wymeersch, B. Denis, V. Sciancalepore, R. D'Errico, A. Clemente, D.-T. Phan-Huy, E. De Carvalho, and P. Popovski, "Reconfigurable, intelligent, and sustainable wireless environments for 6G smart connectivity," *IEEE Commun. Mag.*, vol. 59, no. 10, pp. 99–105, Oct. 2021.

[11] N. Shlezinger, G. C. Alexandropoulos, M. F. Imani, Y. C. Eldar, and D. R. Smith, "Dynamic metasurface antennas for 6G extreme massive MIMO communications," *IEEE Wireless Commun.*, vol. 28, no. 2, pp. 106—-113, Apr. 2021.

[12] Y. Liu, W. Yi, Z. Ding, X. Liu, O. A. Dobre, and N. Al-Dhahir, "Developing NOMA to next generation multiple access: Future vision and research opportunities," *IEEE Wireless Commun.*, vol. 29, no. 6, pp. 120–127, 2022.

[13] S. C. Lam and X. N. Tran, "Improving performance of the typical user in the indoor cooperative NOMA millimeter wave networks with presence of walls," *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, vol. 11, no. 2, pp. e4–e4, 2024.

[14] D.-B. Ha, V.-T. Truong, and Y. Lee, "Performance analysis for rf energy harvesting mobile edge computing networks with SIMO/MISO-NOMA schemes," *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, vol. 8, no. 27, p. e2, 2021.

[15] A. Masaracchia, V.-L. Nguyen, and M. T. Nguyen, "The impact of user mobility into non-orthogonal multiple access (NOMA) transmission systems," *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, vol. 7, no. 24, Oct. 2020.

[16] A. Masaracchia, M. T. Nguyen, and A. Kortun, "User mobility into NOMA assisted communication: Analysis and a reinforcement learning with neural network based approach," *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, vol. 7, no. 25, p. e5, Jan. 2021.

[17] A. Masaracchia, D. B. D. Costa, T. Q. Duong, M.-N. Nguyen, and M. T. Nguyen, "A PSO-based approach for user-pairing schemes in NOMA systems: Theory and applications," *IEEE Access*, vol. 7, pp. 90 550–90 564, 2019.

[18] X. Zhang, C. Han, and X. Wang, "Joint beamforming-power-bandwidth allocation in terahertz noma networks," in *2019 16th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, 2019, pp. 1–9.

[19] H. Zhang, Y. Duan, K. Long, and V. C. M. Leung, "Energy efficient resource allocation in terahertz downlink noma systems," *IEEE Trans. Commun.*, vol. 69, no. 2, pp. 1375–1384, 2021.

[20] Y. Pan, K. Wang, C. Pan, H. Zhu, and J. Wang, "Sum-rate maximization for intelligent reflecting surface assisted terahertz communications," *IEEE Trans. Veh. Technol.*, vol. 71, no. 3, pp. 3320–3325, 2022.

[21] B. Ning, Z. Chen, W. Chen, Y. Du, and J. Fang, "Terahertz multi-user massive MIMO with intelligent reflecting surface: Beam training and hybrid beamforming," *IEEE Trans. Veh. Technol.*, vol. 70, no. 2, pp. 1376–1393, 2021.

[22] F. Zhao, W. Hao, X. You, Y. Wang, Z. Chu, and P. Xiao, "Joint beamforming optimization for IRS-aided thz communication with time delays," *IEEE Wireless Commun. Lett.*, vol. 13, no. 1, pp. 49–53, 2024.
[23] J. Qiao, C. Zhang, A. Dong, J. Bian, and M.-S. Alouini, "Securing intelligent reflecting surface assisted terahertz systems," *IEEE Trans. Veh. Technol.*, vol. 71, no. 8, pp. 8519–8533, 2022.
[24] T. Hou, Y. Liu, Z. Song, X. Sun, Y. Chen, and L. Hanzo, "Reconfigurable intelligent surface aided NOMA networks," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2575–2588, 2020.
[25] B. Zheng, Q. Wu, and R. Zhang, "Intelligent reflecting surface-assisted multiple access with user pairing: NOMA or OMA?" *IEEE Commun. Lett.*, vol. 24, no. 4, pp. 753–757, 2020.
[26] D.-B. Ha, V.-T. Truong, and Y. Lee, "Intelligent reflecting surface assisted rf energy harvesting mobile edge computing NOMA networks: Performance analysis and optimization," *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, vol. 9, no. 32, 2022.
[27] F. Fang, Y. Xu, Q.-V. Pham, and Z. Ding, "Energy-efficient design of IRS-NOMA networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 14 088–14 092, 2020.
[28] J. Zhu, Y. Huang, J. Wang, K. Navaie, and Z. Ding, "Power efficient IRS-assisted NOMA," *IEEE Trans. Commun.*, vol. 69, no. 2, pp. 900–913, 2021.
[29] X. Xu, Q. Chen, X. Mu, Y. Liu, and H. Jiang, "Graph-embedded multi-agent learning for smart reconfigurable THz MIMO-NOMA networks," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, pp. 259–275, 2022.
[30] Y. Pan, K. Wang, and C. Pan, "Intelligent reflecting surfaces-supported terahertz NOMA communications," in *2022 IEEE Wireless Commun. Networking Conf. (WCNC)*, 2022, pp. 1743–1748.
[31] W. Luo, W. Liang, S. Wen, D. Wang, and L. Li, "Power allocation and passive beamforming design for ris aided noma in thz communications," in *2022 10th Int. Conf. Inf. Syst. Computing Technol. (ISCTech)*, 2022, pp. 7–13.
[32] M. H. Kumar, S. Sharma, K. Deka, and M. K. Sharma, "RIS-assisted user pairing noma system for thz communications," in *2023 National Conf. Commun. (NCC)*, 2023, pp. 1–6.
[33] J. Xu, Z. Zhu, Z. Chu, H. Niu, P. Xiao, and I. Lee, "Sum secrecy rate maximization for IRS-aided multi-cluster MIMO-NOMA terahertz systems," *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 4463–4474, 2023.
[34] T. M. Hoang, T. Van Luong, D. Liu, and L. Hanzo, "Ris-aided aanets: Security maximization relying on unsupervised projection-based neural networks," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 2, pp. 2214–2219, 2022.
[35] W. Liang, Z. Ding, Y. Li, and L. Song, "User pairing for downlink non-orthogonal multiple access networks using matching algorithm," *IEEE Transactions on Communications*, vol. 65, no. 12, pp. 5319–5332, 2017.
[36] J. Sayehvand and H. Tabassum, "Interference and coverage analysis in coexisting rf and dense terahertz wireless networks," *IEEE Wireless Communications Letters*, vol. 9, no. 10, pp. 1738–1742, 2020.
[37] H. Zhang, H. Zhang, W. Liu, K. Long, J. Dong, and V. C. M. Leung, "Energy efficient user clustering, hybrid precoding and power optimization in terahertz MIMO-NOMA systems," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 9, pp. 2074–2085, 2020.
[38] J. M. Jornet and I. F. Akyildiz, "Channel modeling and capacity analysis for electromagnetic wireless nanonetworks in the terahertz band," *IEEE Transactions on Wireless Communications*, vol. 10, no. 10, pp. 3211–3221, 2011.
[39] 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, Mar. 2020.
[40] X. Xie, F. Fang, and Z. Ding, "Joint optimization of beamforming, phase-shifting and power allocation in a multi-cluster IRS-NOMA network," *IEEE Trans. Veh. Technol.*, vol. 70, no. 8, pp. 1–1, Aug. 2021.
[41] B. Xia, J. Wang, K. Xiao, Y. Gao, Y. Yao, and S. Ma, "Outage performance analysis for the advanced sic receiver in wireless noma systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 6711–6715, 2018.
[42] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5g systems with randomly deployed users," *IEEE Signal Processing Letters*, vol. 21, no. 12, pp. 1501–1505, 2014.
[43] L. D. Nguyen, H. D. Tuan, T. Q. Duong, O. A. Dobre, and H. V. Poor, "Downlink beamforming for energy-efficient heterogeneous networks with massive MIMO and small cells," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3386–3400, May 2018.
[44] T. Do-Duy, L. D. Nguyen, T. Q. Duong, S. R. Khosravirad, and H. Claussen, "Joint optimisation of real-time deployment and resource allocation for UAV-aided disaster emergency communications," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 11, pp. 3411–3424, 2021.
[45] M. Grant and S. Boyd, "CVX: MATLAB software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.
[46] D. Peaucelle, D. Henrion, and Y. Labit, "Users guide for SeDuMi interface 1.03," 2002. [Online]. Available: http://homepages.laas.fr/peaucell/software/sdmguide.pdf
[47] 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.

Tan Do-Duy received the B.S. degree from the Ho Chi Minh City University of Technology (HCMUT), Vietnam, in 2010, the M.S. degree from the Kumoh National Institute of Technology, South Korea, in 2013, and the Ph.D. degree from the Autonomous University of Barcelona, Spain, in 2019. He is currently with the Department of Computer and Communication Engineering, Ho Chi Minh City University of Technology and Education (HCMUTE), Vietnam, as an Assistant Professor. His main research interests include wireless cooperative communications, real-time optimisation for resource allocation in wireless networks, and coding applications for wireless communications.

Antonino MAsaracchia (Senior Member, IEEE) is currently a Lecturer at the Queen Mary University of London. He was a Research Fellow with Queens University Belfast, U.K. His research interests include 6G networks, digital twin, generative AI and applied machine learning techniques to wireless communications, reconfigurable intelligent surfaces (RIS), UAVenabled networks, and ultra-reliable low-latency communications (URLLC).

Berk Canberk (Senior Member, IEEE) has been an Adjunct Professor with the Department of Electrical and Computer Engineering, Northeastern University, since 2016. He is currently a Professor at Edinburgh Napier University, UK and is also with the Department Artificial Intelligence and Data Engineering, Istanbul Technical University. His current research interests include AI driven network automation and management, software-defined networking, 5G, 6G, and intelligent aerial networks. He was a recipient of the IEEE Turkey Research Incentive Award in 2018, the IEEE INFOCOM Best Paper Award in 2018, the British Council (U.K.) Researcher Link Award in 2017, the IEEE CAMAD Best Paper Award in 2016, and the IEEE INFOCOM Best Poster Paper Award in 2015.

Long D. Nguyen (Member, IEEE) received his Ph.D. degree in Electronics and Electrical Engineering from Queen's University Belfast (QUB), UK, in 2018. He is currently with Duy Tan University in Vietnam as an adjunct Assistant Professor and with Dong Nai University as an Assistant Professor. His research interests include convex optimisation theory and applications for modern wireless systems, resource allocation optimisation, and real-time optimisation and machine learning for embedded systems, wireless communications and Internet of Things (IoTs).

Dr. Nguyen is currently serving as a reviewer for some IEEE Transactions, international journals and conferences. He is a guest editor for EAI Transactions on Industrial Networks and Intelligent Systems. He was awarded the Best Paper Award at the IEEE Digital Signal Processing (DSP) 2017, the IEEE International Conference on Recent Advances in Signal Processing, Telecommunication and Computing (Sigtelcom) 2018, the IEEE International Conference on Communications (ICC) 2019, the International Wireless Communications & Mobile Computing Conference (IWCMC) 2019 and the IEEE Global Communications Conference (GLOBECOM) 2019, 2022. He was also awarded the Exemplary Reviewer Award in IEEE Communications Letters 2018

Trung Q. Duong (Fellow, IEEE) is a Canada Excellence Research Chair (CERC) and a Full Professor at Memorial University of Newfoundland, Canada. He is also the adjunct Chair Professor in Telecommunications at Queen's University Belfast, UK. His current research interests include quantum communications, wireless communications, signal processing, machine learning, and realtime optimisation.

Dr. Duong has served as an Editor/Guest Editor for the IEEE Transactions on Wireless Communications, IEEE Transactions on Communications, IEEE Transactions on Vehicular Technology, IEEE Communications Letters, IEEE Wireless Communications Letters, IEEE Wireless Communications, IEEE Communications Magazines, and IEEE Journal on Selected Areas in Communications. He received the Best Paper Award at the IEEE VTC-Spring 2013, IEEE ICC 2014, IEEE GLOBECOM 2016, 2019, 2022, IEEE DSP 2017, IWCMC 2019, 2023, and IEEE CA- MAD 2023. He has received the two prestigious awards from the Royal Academy of Engineering (RAEng): RAEng Research Chair (2021-2025) and the RAEng Research Fellow (2015-2020). He is the recipient of the prestigious Newton Prize 2017.