IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 15 November 2023, accepted 4 December 2023, date of publication 12 December 2023,
date of current version 20 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3341973

RESEARCH ARTICLE

Performance of Large Intelligent Surfaces in Multiuser Millimeter Wave MIMO-NOMA Systems

NOURAN ARAFAT[1], ENGY ALY MAHER[1], (Senior Member, IEEE), AHMED EL-MAHDY[1], (Senior Member, IEEE), AND FALKO DRESSLER[2], (Fellow, IEEE)

[1]Faculty of Information Engineering and Technology, German University in Cairo, Cairo 16482, Egypt
[2]School of Electrical Engineering and Computer Science, TU Berlin, 10623 Berlin, Germany

Corresponding author: Nouran Arafat (nouran.zaghlool@guc.edu.eg)

This work was supported in part by the German Academic Exchange Service (DAAD); and in part by the Federal Ministry of Education and Research (BMBF), Germany, within the 6G Research and Innovation Cluster (6G-RIC) in co-operation between Technical University (TU) Berlin and the German University in Cairo under Grant 16KISK020K.

ABSTRACT Large Intelligent Surfaces (LIS) are considered a candidate technology for 6G networks due to its simple deployment and relatively low cost. Recently, Millimeter Wave (mmWave) communication has attracted much attention due to its huge spectrum resources (30-300 GHz), which supports gigabits per second data rates. However, it suffers from severe path loss compared to path attenuation at low frequency bands. Massive MIMO compensates this path loss by adding more antennas in a small area. Non-Orthogonal Multiple Access (NOMA) is one of the multiple access techniques that increases the data rate of the system. In this paper, we integrate LIS with multi-user MIMO-NOMA in the mmWave band to enhance the overall system performance. Hybrid precoding is performed by having a zero-forcing digital precoder at the base station and an analog precoding at the LIS. Users are grouped into clusters to minimize multi-user interference. Joint optimization of power allocation of NOMA users, LIS phases, and gains is proposed to maximize the energy efficiency of the system. Closed form expressions for the optimum phases and gains of LIS are obtained using the Quadratic Constraint Quadratic Problem (QCQP). Due to the non-convexity of the formulated problem, alternating optimization is developed to solve the problem. Subsequently, the optimization parameters are decoupled through applying fractional programming. The problem is divided into sub problems solved alternatively. The performance of the system is evaluated in terms of energy efficiency and power consumption. Numerical results show the superior performance of LIS at the expense of higher power consumption.

INDEX TERMS Large intelligent surface (LIS), millimeter wave (mmWave) communication, NOMA, MIMO, passive RIS, active RIS, spectral efficiency, energy efficiency, joint optimization.

GLOSSARY

ADC	Analog Digital Converter.
AF	Amplify and Forward.
BS	Base Station.
DAC	Digital Analog Converter.
DoF	Degree of Freedom.
FDD	Frequency Division Duplexing.
FDMA	Frequency Division Multiple Access.
FP	Fractional Programming.
LIS	Large Intelligent Surface.
NOMA	Non Orthogonal Multiple Access.
OMA	Orthogonal Multiple Access.
QCQP	Quadratic Constraint Quadratic Problem.
RE	Reflecting Element.
RF	Radio Frequency.
RIS	Reconfigurable Intelligent Surface.
SC	Superposition Coding.

The associate editor coordinating the review of this manuscript and approving it for publication was Mehdi Sookhak.

SIC	Successive Interference Cancellation.
TDMA	Time Division Multiple Access.

I. INTRODUCTION

Nowadays, Millimeter Wave (mmWave) communication has much attention as it supports data rates of gigabits per second due to its huge spectrum resources (30-300 GHz). However, it experiences severe path loss compared to lower frequency bands [1], [2]. Integrating massive MIMO with mmWave compensates the path loss by adding more antennas in a small area [3]. The key idea of massive MIMO is to deploy a large number of antennas N to serve M users simultaneously.

Non-Orthogonal Multiple Access (NOMA) is one of the most promising technologies, due to its high spectral and energy efficiencies [4]. In the same resource block, multiple users are served using power-domain NOMA. By using superposition coding (SC) at the transmitter and successive interference cancellation (SIC) at the receiver, NOMA improves the spectral efficiency of the wireless system.

Reconfigurable Intelligent Surface (RIS) has raised significant interest in the current research, because it is relatively cheap and easy to mount on building facades. A RIS is mainly a two dimensional meta-surface with the ability to control the wireless signal and reflect it in a desired direction. In other words, RIS consists of passive reflecting elements (REs). Each RE has the capability to reflect the incident electromagnetic waves using the reflecting coefficient (amplitude and phase) [5]. By adjusting the reflecting coefficients, the reflected signal is added constructively in the desired direction (receiver) and destructively in other directions (suppressing interference). Unlike other related technologies, such as relays and MIMO beamforming, RIS does not require its own power source [6]. Thus, RIS can enhance the spectral and energy efficiency of the wireless system. By deploying RIS in mmWave MIMO system, the performance of the channel is improved.

Based on power consumption, there are mainly two types of RIS: active and passive. Active RIS is sometimes referred to in the literature as Large Intelligent Surface (LIS), where each RE is connected to a power amplifier and phase shifter. On the other hand, passive RIS, or reflecting RIS, consists of passive elements with almost zero power consumption. Each element has a different phase shifter value to direct the signal in a certain direction [6]. A comparison between RIS and relay is addressed in [7] and [8]. Furthermore, LIS (active RIS) differs from amplify and forward relay (AF) in terms of hardware complexity and power consumption. The AF relay amplifies and retransmits the signal, thus requiring two time slots. Additionally, relays require RF chains, analog to digital converters (ADC), digital to analog converters (DAC) and amplifiers [8].

The motivation behind integrating NOMA with LIS is to combine the benefits of both technologies into one system, enhancing spectral efficiency, user coverage and connectivity at a low cost. Since the wireless channel suffers from high fluctuations due to path loss and blockage, the combination of these technologies can control the wireless channel and guarantee specific users' order required by NOMA.

In this paper, we propose a joint iterative optimization algorithm for LIS phase shifts and gains, as well as power allocation for users, to maximize energy efficiency. The main contributions of this paper can be summarized as follows:

1) An algorithm for clustering the users is proposed to minimize the multi-user interference such that users in the same cluster have a high channel correlation, while users in different clusters have a low correlation between their channels.
2) Hybrid precoding is performed by zero-forcing digital precoding at the BS and analog precoding at LIS. LIS elements are divided into sub-surfaces and each subsurface is assigned to one antenna at the BS to realize an equivalent performance as the energy-efficient sub-connected architecture.
3) A joint optimization of power allocation among users, LIS power amplifier gain, and LIS phase shifts are proposed for the mmWave MIMO-NOMA multi-user system.

The rest of this paper is organized as follows. Section II introduces the related work. Section III presents the system and channel model. Then, Section IV shows the signal model and signal to interference noise ratio for the mmWave MIMO-NOMA aided LIS system. Afterwards, the energy efficiency maximization problem is given and transformed into simpler form in Section V. Section VI derives the optimum values for LIS parameters (phase and gain) and users power. Then, the numerical results are shown in Section VII. Finally, the conclusions are presented in Section VIII.

Notation: Upper and lower bold case represent matrix and vector respectively. The notation of $(.)^T$,$(.)^H$,$(.)^{-1}$ are the transpose, conjugate transpose (Hermitian) and the inverse of the matrix respectively. The $||.||_p$ represents the l_p norm. If S is a set, then $|S|$ denotes the number of elements inside it. The difference between two sets is denoted by $(\setminus)$.The Kronecker product is denoted by $\otimes$. The dot product between two vectors is denoted by $\odot$. The notation of $CN(\mu, \sigma^2)$ denotes the complex Gaussian distribution with mean μ and covariance σ^2. The expectation of random variable is denoted by $E\{.\}$.

II. RELATED WORK

The concept of RIS has been, for example, summarized in [9]. The continuous and discrete RIS phases are explained. Integrating RIS into a two-user Non-Orthogonal Multiple Access (NOMA) system is discussed. Moreover, the working principles and challenges of RIS-aided systems are presented in [10].

In [11], Intelligent Reflecting Surface (IRS) is proposed to support a multi-user millimeter wave (mmWave) system, in which multiple IRS are deployed to enhance the coverage

of the system when the direct links between the base station and users are blocked. The authors present an optimization for the weighted sum rate maximization problem for joint active and passive beamforming. To address this issue, an alternating iterative algorithm is proposed to decouple the active and passive beamforming matrices. In [12], multiple IRS are deployed in multi-user system to enhance the overall system performance. Double-manifold alternating optimization is proposed to maximize the weighted sum rate and optimize the beamforming matrix at the base station (BS) and the reflection coefficients of each IRS element. Results show that the proposed approach outperforms traditional methods in terms of spectral efficiency. In [13], RIS assisted aerial-ground communication scenario, exploring UAV trajectory design and RIS' phase shift optimization aiming to maximize the spectral efficiency. The *Riemannian conjugate gradient method* is used to solve the unit modulus constraint of the reflecting coefficient of RIS.

RIS aided frequency division duplexing (FDD) system for uplink and downlink transmission is discussed in [14]. The same phase shift matrix for both transmission directions is proposed. RIS is used as interference nulling in a multi-user environment [15]. In [16], joint optimization of digital precoder at BS and the phase shift at RIS is considered in order to enhance the received signal power.

There are many reports demonstrating the performance of passive RIS; however, active RIS is only addressed in a few works [17] and [18]. A comparison between active and passive RIS is investigated in [18] and [19]. In [17], two architectures for active RIS are compared: fully connected and sub-connected architectures. In fully connected architecture, each RIS element has one power amplifier. However, in sub-connected architecture, a group of RIS elements share the same power amplifier. As the number of RIS elements increases, the fully connected architecture consumes huge amount of power consumption. Therefore, the sub-connected architecture reduces the power consumption at the expense of lower degree of freedoms (DoF) for beamforming design. The spectral and energy efficiencies are evaluated, and the results show that the latter architecture enjoys higher energy efficiency at the cost of lower sum rate than fully connected architecture. In [18], active and passive RIS are considered in case of a strong/weak direct link. Results show that passive RIS has only 3% gain compared to the case without RIS in case of a strong direct link, while active RIS achieves noticeable spectral efficiency gain of 67% in case of a strong direct link. Thus, by implementing active RIS, multiplicative fading can be overcome.

Active and passive RIS aided systems are compared in [19]. For fair comparison, the same overall power budget is assumed at the base station (BS) and RIS. The system is evaluated based on the achievable sum rate. The results show that active RIS has superior performance when the power budget is not small and the number of RIS elements is small. This is because as the number of RIS elements increases, a huge amount of power is consumed to supply the power amplifiers, and the thermal noise also increases in active RIS. In [20], conventional hybrid precoder is replaced by an energy-efficient RIS to realize analog precoder. Also, a deep learning-based multiple discrete classification is proposed to solve the sum rate maximization optimization problem.

Non-Orthogonal Multiple Access (NOMA) is widely demanded due to the increasing number of devices and enormous connectivity [21]. The combination of NOMA with IRS is addressed in [22], [23], [24], [25], [26], and [27]. In [22], the role of RIS in MIMO-NOMA is discussed in different cases. Different Multiple Access (MA) techniques are investigated with RIS, i.e., FDMA, TDMA, and NOMA [23]. Results reveal that integrating RIS with NOMA system enhances the spectral efficiency. In [24], the spectral efficiency optimization is only addressed in case of passive RIS. DC programming and successive convex approximation (SCA) are utilized to solve EE maximization problem in NOMA-RIS system [25]. However, authors of [23], [24], and [25] did not consider active RIS. Also, mmWave communications did not consider in [23] and [25]. In [28], authors proposed two low complexity ZF approximation schemes for precoding and user grouping. Those schemes are based on dividing the antennas at BS into sub-arrays and grouping users according to their elevation angles. A plane wave approximation is considered in which a precoding filters are designed to achieve the performance of ZF precoder with lower complexity. While in our paper, user grouping is based on the Euclidean distance, where a cluster head users are chosen for each cluster. The remaining users are assigned to the nearest cluster based on the Euclidean distance between the candidate user and the cluster head users. Zero forcing precoding is adopted to illuminate the inter-cluster interference. In [29], a comparison between different precoding techniques (zero-forcing (ZF), maximum ratio transmission/ combining (MRT/MRC), and minimum mean squared error (MMSE)) are discussed for energy efficiency multi-user MIMO system. Results show that the highest EE is achieved using ZF precoding due to interference suppression at affordable complexity.

A RIS partitioning technique is proposed in [26], in which each sub-surface serves one user and only a single antenna is considered at the BS. In [27], only two users NOMA aided IRS system are considered. A comparison between IRS-OMA and IRS-NOMA is presented. Results show superior energy efficiency performance in the case of NOMA. Both [26] and [27] do not consider mmWave band and consider passive RIS. To the best of our knowledge, no paper integrates NOMA with LIS in mmWave band, which is considered in this paper.

Based on the discussed related work, we believe that integrating MIMO-NOMA with LIS in mmWave band will enhance the overall system performance. In this paper, the

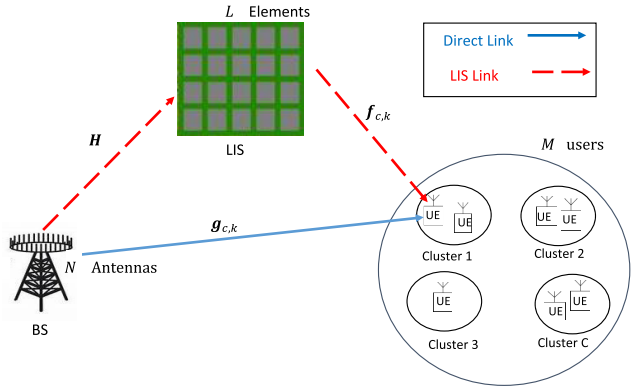

FIGURE 1. System model of downlink mmWave MIMO-NOMA aided LIS.

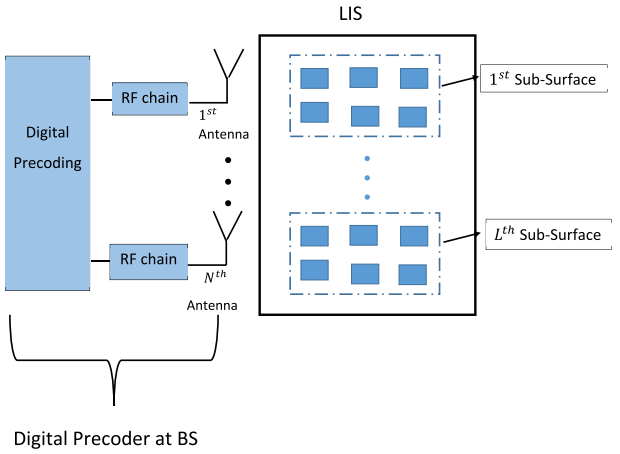

FIGURE 2. Hybrid precoder with integrating LIS as analog precoder.

three technologies are integrated for increasing the spectral efficiency, extending the coverage and minimizing the cost. Combining MIMO-NOMA with LIS in mmWave band is still missing in the literature. In addition, users grouping algorithm and hybrid precoding are proposed. The hybrid precoding is implemented by considering digital precoding at the BS, and LIS acts as an analog precoder.

III. SYSTEM MODEL

In this section, the system model and the channel model are presented. We consider a downlink mmWave MIMO-NOMA system assisted by LIS, where the base station is equipped with N antenna elements, and LIS has L elements to serve M single antenna users, as shown in Fig. 1. The base station transmits data to users through the direct link and the reflected link from LIS. Users are grouped into clusters to minimize the interference among them. The channel between BS and k^{th} user in the c^{th} cluster, the channel between BS and LIS, and the channel between LIS and k^{th} user in the c^{th} cluster are represented by $\boldsymbol{g}_{c,k} \in \mathbb{C}^{N\times 1}, \boldsymbol{H} \in \mathbb{C}^{N\times L}, \boldsymbol{f}_{c,k} \in \mathbb{C}^{L\times 1}$.

A. CHANNEL MODEL

The mmWave channel model for the k^{th} user in c^{th} cluster is given as [3]

$$\boldsymbol{g}_{c,k} = \sqrt{\frac{N}{P_{c,k}}} \sum_{p=1}^{P_{c,k}} \alpha_{c,k}^{(p)} \mathbf{a}(\varphi_{c,k}^{(p)}, \theta_{c,k}^{(p)}), \tag{1}$$

where $P_{c,k}$ is the number of paths for the k^{th} user in the c^{th} cluster group. The complex gain of the p^{th} path is given by $\alpha_{c,k}^{(p)}$. The azmiuth angle of departure (AOD) and the elevation angle of arrival (AOA) are denoted by $\varphi_{c,k}^{(p)}, \theta_{c,k}^{(p)}$ respectively. The steering vector is given by $\mathbf{a}(\varphi_{c,k}^{(p)}, \theta_{c,k}^{(p)})$ with size $N \times 1$. A uniform planar array (UPA) is considered with N_1 antenna elements in horizontal direction and N_2 antennas in vertical direction where $N = N_1N_2$, so $\mathbf{a}(\varphi, \theta) = \mathbf{a}_{azmi}(\varphi) \otimes \mathbf{a}_{elev}(\theta)$, where $\mathbf{a}_{azmi}(\varphi) = \frac{1}{\sqrt{N_1}}[1,\ e^{j2\pi(\frac{d_1}{\lambda})sin(\varphi)}, .., e^{j2\pi(N_1-1)(\frac{d_1}{\lambda})sin(\varphi)}]$, and $\mathbf{a}_{elev}(\theta) = \frac{1}{\sqrt{N_2}}[1,\ e^{j2\pi(\frac{d_2}{\lambda})sin(\theta)}, .., e^{j2\pi(N_2-1)(\frac{d_2}{\lambda})sin(\theta)}]$, λ is the signal wavelength and d_1 and d_2 are the antenna spacing in horizontal and vertical directions respectively, with $d_1 = d_2 = \lambda/2$.

B. USER CLUSTERING AND HYBRID PRECODING

Users are grouped into clusters, where one cluster head user is chosen for each cluster. Cluster head users are selected based on the maximum Euclidean distance between each other. First, the distance between all users are calculated. Then the two users with the maximum distance between each other are considered to be the heads of the first two clusters. In order to obtain the next cluster head, the distances between the remaining users and the two chosen cluster heads are calculated. If the distances between any user and the first two cluster heads exceed certain threshold, this user is considered as the head of the next cluster, and so on. If there is no user with distance more than the threshold, then the threshold is reduced by 10%. Afterwards, the remaining users are assigned to the closest cluster. This leads to users in the same cluster enjoy a high channels correlation, while users in different clusters have low correlation.

Hybrid precoding is performed by having a digital precoder at the BS and analog precoder at LIS. An equivalent performance of energy efficient sub-connected architecture is achieved by dividing the LIS elements equally over S sub-surfaces as shown in Fig. 2. Each RF chain is only corresponding to one sub-surface. The number of sub-surfaces S is equal to the number of antennas N at BS. The number of elements per sub-surface is $L_S = L/N$, where L_S must be integer. The interference between sub-surfaces can be ignored by carefully designed the size of sub-surface and the distance between RF chains [20]. Thus, the RF chain's signal can only be considered from the corresponding s^{th} sub-surface. It is noted that, sub-connected architecture of LIS is not directly implemented, however an equivalent

performance can be achieved by dividing the elements over sub-surfaces in which each subsurface is connected to one antenna at BS. The channel between the BS and the LIS elements is given by

$$H=\begin{bmatrix} \underbrace{\hbar_{1,1},..,\hbar_{1,L_S}} & 0 & 0 \\ 0 & \underbrace{\hbar_{2,L_S+1},..,\hbar_{2,2L_S}} & 0 \\ 0 & 0 & \underbrace{\hbar_{N,L-L_S+1},..,\hbar_{N,L}} \end{bmatrix}$$

where $\hbar_{n,l}$ is the channel between the n^{th} antenna at BS and the l^{th} LIS element.

The precoding matrix of LIS is denoted by $\boldsymbol{\Phi}$=diag($\boldsymbol{\phi} \odot \boldsymbol{a}$) $\in \mathbb{C}^{L\times L}$. The phase vector of LIS elements is given by $\boldsymbol{\phi} = [\phi_1, \phi_2, \ldots, \phi_L] \in \mathbb{C}^{L\times 1}$, where $\phi_l = e^{j\theta_l}$. The gain vector for LIS elements is denoted by $\boldsymbol{a} = [a_1, a_2, \ldots, a_L] \in \mathbb{R}^{L\times 1}$. In passive RIS, the gain vector is assumed to be all ones ($\mathbf{1}_L$).

IV. MATHEMATICAL FORMULATION

Clustering algorithm is applied in order to group users. The total number of clusters is denoted by C. The set U_i contains the index of users in i^{th} cluster. So, $|U_i|$ represents the total number of users in i^{th} cluster. Every cluster has at least one user. The received signal at LIS is modeled as

$$\boldsymbol{r} = \underbrace{\boldsymbol{\Phi}\boldsymbol{H}^H \sum_{i=1}^{C}\sum_{j=1}^{|U_i|} \boldsymbol{w}_i\sqrt{p_{i,j}}s_{i,j}}_{desired\ signal} + \underbrace{\boldsymbol{\Phi}\boldsymbol{z}}_{dynamic\ noise} + \underbrace{\boldsymbol{n}_a}_{static\ noise} \quad (2)$$

where $s_{i,j}$ is the transmitted signal with $E\{|s_{i,j}|^2\} = 1$ and $p_{i,j}$ is the transmitted power for j^{th} user and i^{th} cluster. The beamforming vector at BS for i^{th} cluster is denoted by $\boldsymbol{w}_i$, where the number of beams from BS is equal to the number of clusters C. The noise introduced due to active RIS can be divided into two parts: dynamic and static noise. The dynamic noise is denoted by ($\boldsymbol{\Phi}\ z$), while the static noise is denoted by $\boldsymbol{n}_a$. The dynamic noise is introduced due to the active hardware components, i.e., reflection type amplifier, which amplifies the incident signal as well as the noise. This noise depends on the input power applied to the power amplifier circuit. It also depends on the range in which the power amplifier is operating. The dynamic noise has a significant value and cannot be neglected in the analysis. Where z is as result of the input noise and the inherent device noise of active elements [30]. On the other hand, the static noise is because of the phase shifter circuit and it does not depend on $\boldsymbol{\Phi}$. In consequence, static noise is usually neglected when compared to dynamic noise. The received signal $y_{c,k}$ for k^{th} user in c^{th} cluster is given as

$$\begin{aligned} y_{c,k} &= \boldsymbol{h}_{c,k}^H \sum_{i=1}^{C}\sum_{j=1}^{|U_i|} \boldsymbol{w}_i\sqrt{p_{i,j}}s_{i,j} + \boldsymbol{f}_{c,k}^H\boldsymbol{\Phi}\boldsymbol{z} + n_{c,k} \\ &= \underbrace{\boldsymbol{h}_{c,k}^H\boldsymbol{w}_c\sqrt{p_{c,k}}s_{c,k}}_{desired\ signal} \\ &\quad + \underbrace{\boldsymbol{h}_{c,k}^H\boldsymbol{w}_c\left(\sum_{j=1}^{k-1}\sqrt{p_{c,j}}s_{c,j} + \sum_{j=k+1}^{|U_c|}\sqrt{p_{c,j}}s_{c,j}\right)}_{intra-cluster\ interferences} \\ &\quad + \underbrace{\boldsymbol{h}_{c,k}^H\sum_{i\neq c}^{C}\boldsymbol{w}_i\sum_{j=1}^{|U_i|}\sqrt{p_{i,j}}s_{i,j}}_{inter-cluster\ interferences} \\ &\quad + \underbrace{\boldsymbol{f}_{c,k}^H\boldsymbol{\Phi}\boldsymbol{z}}_{noise\ due\ to\ active\ RIS} + \underbrace{n_{c,k}}_{noise}, \end{aligned} \quad (3)$$

where $\boldsymbol{h}_{c,k} = \boldsymbol{g}_{c,k} + \boldsymbol{H}\boldsymbol{\Phi}\boldsymbol{f}_{c,k}$ is the equivalent channel from BS to k^{th} user in the c^{th} cluster, the first term represents the direct link from BS to user, however the second part in $\boldsymbol{h}_{c,k}$ represents the reflected link from LIS. The received signal $y_{c,k}$ at k^{th} user in c^{th} cluster in (3) is divided into: the desired signal, the interference from users in same cluster, interference due to users in other cluster, noise due to integrating LIS in the system and the noise at the receiver, respectively.

Digital Zero-Forcing (ZF) precoding vector is used at the BS and it is denoted by $\boldsymbol{w}_i \in \mathbb{C}^{N\times 1}$ for the i^{th} cluster. The precoding matrix $W \in \mathbb{C}^{N\times C}$ is given by

$$\boldsymbol{W} = [\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_c] = \hat{\boldsymbol{H}}(\hat{\boldsymbol{H}}^H\hat{\boldsymbol{H}})^{-1} \quad (4)$$

where $\hat{\boldsymbol{H}} = [\boldsymbol{h}_{1,\underline{k}}, \boldsymbol{h}_{2,\underline{k}}, \ldots, \boldsymbol{h}_{C,\underline{k}}]$ and $\boldsymbol{h}_{c,\underline{k}}$ represents the k^{th} user with the highest channel gain in the c^{th} cluster.

The dynamic noise follows additive white Gaussian distribution, i.e., $z \sim \mathcal{CN}(\mathbf{0}_L, \sigma_z^2\boldsymbol{I}_L)$ and the static noise of LIS ($\boldsymbol{n}_a$) is neglected. While $n_{c,k}$ represents the noise at the user with $\mathcal{CN}(0, \sigma^2)$ distribution. Assuming users are order in each cluster based on channel gain in descending order. In (3), the second summation in the second term (intra-cluster interference) with summation greater than k is cancelled due to applying NOMA using successive interference cancellation (SIC). Then, the signal to noise and interference ratio (SINR) for k^{th} user in c^{th} cluster is given by

$$\begin{aligned} \gamma_{c,k} &= (|\boldsymbol{h}_{c,k}^H\boldsymbol{w}_c|^2\, p_{c,k}).(|\boldsymbol{h}_{c,k}^H\boldsymbol{w}_c|^2 \sum_{j=1}^{k-1} p_{c,j} \\ &\quad + \sum_{i\neq c}^{C}|\boldsymbol{h}_{c,k}^H\boldsymbol{w}_i|^2 \sum_{j=1}^{|U_i|} p_{i,j} + ||\boldsymbol{f}_{c,k}^H\boldsymbol{\Phi}||^2\sigma_z^2 + \sigma^2)^{-1} \end{aligned} \quad (5)$$

The spectral efficiency of the k^{th} user in the c^{th} cluster is

$$R_{c,k} = log_2(1 + \gamma_{c,k}). \quad (6)$$

The total spectral efficiency is written as

$$R_{sum} = \sum_{c=1}^{C} \sum_{k=1}^{|U_c|} R_{c,k}, \tag{7}$$

The energy efficiency is defined as the ratio of the total spectral efficiency to the total power consumption

$$EE = \frac{R_{sum}}{P_c^{tot}} (bps/Hz/W), \tag{8}$$

where P_c^{tot} is the total power consumption. The power consumption of active RIS is due to the switches, phase shift control circuit and the power amplification circuit. There are many factors that leads to power dissipation in active RIS, i.e., the hardware implementation, DC power, output power as well as load characteristics. While for the passive RIS, the power dissipation is only due to the switches and phase shift control circuit ($P_{passive} = L\,P_{PS}$).

Then, the total power consumption is given as the power dissipated in the BS, active RIS and users as

$$\begin{aligned} P_c^{tot} = & \underbrace{\zeta P_{tx} + P_B}_{\text{power dissipated due to BS}} \\ & + \underbrace{\xi \Big(\sum_{i=1}^{C} \sum_{j=1}^{|U_i|} ||\boldsymbol{\Phi} \boldsymbol{H}^{\boldsymbol{H}} \boldsymbol{w}_i||^2 p_{i,j} + ||\boldsymbol{\Phi}||^2 \sigma_z^2\Big)}_{\text{power dissipated due to active RIS}} \\ & + \underbrace{L(P_{PS} + P_{PA})}_{\text{power dissipated by phase shift and power amplifier}} \\ & + \underbrace{MP_u}_{\text{power dissipated by users}}, \end{aligned} \tag{9}$$

where ζ and ξ are the inverse energy conversion coefficients at the BS and LIS, respectively. In addition, P_B is the base-band power consumption. While P_{tx} is the total transmitted power at BS, i.e., $P_{tx} = \sum_{i=1}^{C} \sum_{j=1}^{|U_i|} ||\boldsymbol{w}_i||^2 p_{i,j}$. The power dissipated per user is denoted by P_u. However the second term in (9) represents the power dissipated due to LIS. P_{PS} and P_{PA} are the static power of LIS hardware for the phase shift circuit and the power amplifier, respectively.

V. FORMULATION OF THE ENERGY EFFICIENCY OPTIMIZATION

The optimization problem is presented in this section. The objective function is to maximize the energy efficiency with respect to the user power ($\boldsymbol{p}$) and LIS phase ($\boldsymbol{\phi}$) and gain ($\boldsymbol{a}$).

The maximization problem is written as

$$\max_{\boldsymbol{p},\boldsymbol{\phi},\boldsymbol{a}} \quad EE \tag{10a}$$

$$\text{s.t.} \quad C_1 : p_{i,j} \geq 0, \quad \forall i \in C, j \in |U_i| \tag{10b}$$

$$C_2 : \zeta \sum_{i=1}^{C} \sum_{j=1}^{|U_i|} ||\boldsymbol{w}_i||^2 p_{i,j} \leq \tilde{P}_{max}^{BS}, \tag{10c}$$

$$C_3 : \xi \Big(\sum_{i=1}^{C} \sum_{j=1}^{|U_i|} ||\boldsymbol{\Phi} \boldsymbol{H}^{\boldsymbol{H}} \boldsymbol{w}_i||^2 p_{i,j} + ||\boldsymbol{\Phi}||^2 \sigma_z^2\Big) \leq \tilde{P}_{max}^{RIS}, \tag{10d}$$

$$C_4 : |\phi_l| = 1, \quad \forall l \in L \tag{10e}$$

$$C_5 : a_l \geq 0, \quad \forall l \in L, \tag{10f}$$

where C_1 is to guarantee that the power assigned to each user must be positive. C_2 indicates that the total power consumption by the BS is less than the maximum power assigned to the BS (P_{max}^{BS}). However C_3 guarantees that the power consumed due to LIS does not exceed the maximum power dedicated for LIS (P_{max}^{RIS}). The last two constraints for the unit modulus phase shift constraint, and power amplifier can be any positive number.

Let us define $\tilde{P}_{max}^{BS} = P_{max}^{BS} - P_B$, $\tilde{P}_{max}^{RIS} = P_{max}^{RIS} - (L(P_{PS} + P_{PA}))$ and $\boldsymbol{p} = [p_{1,1}, p_{1,2}, p_{1,|U_1|}, \ldots . p_{C,|U_C|}]^T$ is a vector of the power allocated for all users $\in \mathbb{R}_+^{M \times 1}$. The optimal solution does not affect by the base of the logarithmic function, thus natural logarithmic is considered through the rest of the paper.

The problem in (10) is non-convex because the Hessian matrix is negative since it is a non-linear function consisting of fractional and logarithmic functions. The non-convexity implies that the problem may has multiple local optima or discontinuities. To solve this non-convex problem, Dinkelbach's algorithm is applied to make the problem in a form of a non-fractional so the Hessian matrix becomes non-negative. Dinkelbach's algorithm is an efficient method to solve fractional problem by decoupling the numerator and denominator of a single ratio [31]. Thus, Dinkelbach's algorithm transformed (10a) into parameterized problem as

$$f_1(\boldsymbol{p}, \boldsymbol{\phi}, \boldsymbol{a}) = R_{sum} - \eta P_c^{tot} \tag{11}$$

where $\eta = \frac{R_{sum}}{P_c^{tot}}$ is an auxiliary variable. The optimal energy efficiency η^{opt} satisfies this condition $f_1(\boldsymbol{p}^*, \boldsymbol{\phi}^*, \boldsymbol{a}^*) = 0$. The auxiliary variable η is updated iteratively until the condition is satisfied. Now, the maximization problem is given as

$$\begin{aligned} \max_{\boldsymbol{p},\boldsymbol{\phi},\boldsymbol{a}} \quad & f_1(\boldsymbol{p}, \boldsymbol{\phi}, \boldsymbol{a}) \\ \text{s.t.} \quad & C_1, C_2, C_3, C_4, C_5 \end{aligned} \tag{12}$$

In order to solve (12), which is still non convex problem. We apply fractional programming. A closed form for fractional programming (FP) was proposed to deal with sum of logarithms of ratio [32]. The SINR in (5), can be

transformed to

$$\gamma_{c,k} = \frac{|A_{c,k}|^2}{B_{c,k} - |A_{c,k}|^2} \tag{13}$$

where

$$A_{c,k} = \boldsymbol{h}_{c,k}^H \boldsymbol{w}_c \sqrt{p_{c,k}} \tag{14a}$$

$$\begin{aligned} B_{c,k} = &\Bigg(\sum_{j=1}^{k} |\boldsymbol{h}_{c,k}^H \boldsymbol{w}_c \sqrt{p_{c,j}}|^2 \\ &+ \sum_{i \neq c}^{C} \sum_{j=1}^{|U_i|} |\boldsymbol{h}_{c,k}^H \boldsymbol{w}_i \sqrt{p_{i,j}}|^2 \\ &+ ||\boldsymbol{f}_{c,k}^H \boldsymbol{\Phi}||^2 \sigma_z^2 + \sigma^2 \Bigg) \end{aligned} \tag{14b}$$

By introducing an auxiliary variable $\mu_{c,k}$, the logarithmic function is tackled based on

$$\begin{aligned} \ln(1+\gamma_{c,k}) = \max_{\mu_{c,k} \geq 0} \ln(1+\mu_{c,k}) - \mu_{c,k} \\ + (1+\mu_{c,k}) \left[\frac{\gamma_{c,k}}{1+\gamma_{c,k}} \right] \end{aligned} \tag{15}$$

From (13), we can write:

$$\frac{\gamma_{c,k}}{1+\gamma_{c,k}} = \frac{|A_{c,k}|^2}{B_{c,k}}$$

After substituting the previous expression in (15), the problem is written as

$$\begin{aligned} \max_{\boldsymbol{p},\boldsymbol{\phi},\boldsymbol{a},\boldsymbol{\mu}} \quad & \sum_{c=1}^{C} \sum_{k=1}^{|U_c|} \ln(1+\mu_{c,k}) - \mu_{c,k} + (1+\mu_{c,k}) \frac{|A_{c,k}|^2}{B_{c,k}} \\ \text{s.t.} \quad & C_1, C_2, C_3, C_4, C_5, \mu_{c,k} \geq 0 \end{aligned} \tag{16}$$

where $\boldsymbol{\mu} = [\mu_{1,1}, \mu_{1,2}, .., \mu_{1,|U_1|}, \ldots, \mu_{C,|U_C|}]^T \in \mathbb{C}^{M \times 1}$. For a given $\boldsymbol{\mu}$, one can focus on the sum of ratio problem as

$$\max_{\boldsymbol{p},\boldsymbol{\phi},\boldsymbol{a},\boldsymbol{\mu}} \quad \sum_{c=1}^{C} \sum_{k=1}^{|U_c|} \frac{|A_{c,k}|^2}{B_{c,k}} \tag{17}$$

By introducing new auxiliary variable $\boldsymbol{\nu} = [\nu_{1,1}, ., \nu_{C,|U_C|}] \in \mathbb{C}^{M \times 1}$, using quadratic transform, problem (17) is equivalent to

$$\max_{\boldsymbol{p},\boldsymbol{\phi},\boldsymbol{a},\boldsymbol{\mu}\boldsymbol{\nu}} \quad \sum_{c=1}^{C} \sum_{k=1}^{|U_c|} 2\mathfrak{Re}\{\nu_{c,k}^* A_{c,k}\} - |\nu_{c,k}|^2 B_{c,k} \tag{18}$$

After substituting by (18), the original optimization problem (12) is given by

$$\begin{aligned} \max_{\boldsymbol{p},\boldsymbol{\phi},\boldsymbol{a},\boldsymbol{\mu}\boldsymbol{\nu}} \quad & f_2(\boldsymbol{p},\boldsymbol{\phi},\boldsymbol{a},\boldsymbol{\mu},\boldsymbol{\nu}) \\ \text{s.t.} \quad & C_1, C_2, C_3, C_4, C_5 \end{aligned} \tag{19}$$

where

$$\begin{aligned} f_2(\boldsymbol{p},\boldsymbol{\phi},\boldsymbol{a},\boldsymbol{\mu},\boldsymbol{\nu}) = &\sum_{c=1}^{C} \sum_{k=1}^{|U_c|} [\ln(1+\mu_{c,k}) - \mu_{c,k}] \\ &+ \sum_{c=1}^{C} \sum_{k=1}^{|U_c|} 2\sqrt{1+\mu_{c,k}} \mathfrak{Re}\{\nu_{c,k}^* A_{c,k}\} \\ &- \sum_{c=1}^{C} \sum_{k=1}^{|U_c|} |\nu_{c,k}|^2 B_{c,k} - \eta P_c^{tot} \end{aligned}$$

The optimal solution of this problem is obtained using alternating joint optimization [33]. The main idea is to optimize each variable independently while assuming the other variables are fixed.

VI. JOINT OPTIMIZATION OF USER POWER, LIS PHASE AND GAIN

In this section, problem (19) is solved using alternating optimization. The objective is to obtain the optimum value for the user power allocation ($\boldsymbol{p}^{opt}$), the auxiliary variables ($\boldsymbol{\mu}^{opt}$ & $\boldsymbol{\nu}^{opt}$) and LIS parameters ($\boldsymbol{\phi}^{opt}$ & $\boldsymbol{a}^{opt}$).

A. OPTIMIZATION OF POWER ALLOCATION

In this section, the optimum user power is derived. Assume the other parameters ($\boldsymbol{\mu}$, $\boldsymbol{\nu}$,$\boldsymbol{\phi}$, $\boldsymbol{a}$) are fixed. The optimization problem (19) can be rewritten as

$$\max_{\boldsymbol{p}} \quad f_3(\boldsymbol{p}) \tag{20a}$$

$$\text{s.t.} \quad C_1 : p_{i,j} \geq 0, \quad \forall i \in C, j \in |U_i| \tag{20b}$$

$$C_2 : \sum_{i=1}^{C} \sum_{j=1}^{|U_i|} ||\boldsymbol{w}_i||^2 p_{i,j} \leq \bar{P}_{max}^{BS}, \tag{20c}$$

where

$$\begin{aligned} f_3(\boldsymbol{p}) = &\sum_{c=1}^{C} \sum_{k=1}^{|U_c|} \bar{A}_{c,k} \mathfrak{Re}\{\bar{B}_{c,k} \sqrt{p_{c,k}}\} \\ &- \sum_{c=1}^{C} \sum_{k=1}^{|U_c|} |\nu_{c,k}|^2 \Bigg[\sum_{j=1}^{k} |D_{c,k,c} \sqrt{p_{c,j}}|^2 \\ &+ \sum_{i \neq c}^{C} \sum_{j=1}^{|U_i|} |D_{c,k,i} \sqrt{p_{i,j}}|^2 + E_{c,k} \Bigg] \\ &- \eta\zeta \left(\sum_{i=1}^{C} \sum_{j=1}^{|U_i|} ||\boldsymbol{w}_i||^2 p_{i,j} \right) - Q \\ &- \eta\xi \left(\sum_{i=1}^{C} \sum_{j=1}^{|U_i|} Z_i p_{i,j} + Y \right) + X \end{aligned}$$

where

$$\bar{A}_{c,k} = 2\sqrt{1+\mu_{c,k}}, \quad D_{c,k,i} = \boldsymbol{h}_{c,k}^H \boldsymbol{w}_i, \quad \bar{B}_{c,k} = v_{c,k}^* D_{c,k,c},$$
$$E_{c,k} = ||\boldsymbol{f}_{c,k}^H \boldsymbol{\Phi}||^2 \sigma_z^2 + \sigma^2, \qquad \bar{P}_{max}^{BS} = \zeta^{-1} \tilde{P}_{max}^{BS},$$
$$Q = \eta\Big[P_B + MP_u + L(P_{PS} + P_{PA})\Big], \quad Z_i = ||\boldsymbol{\Phi}\boldsymbol{H}^{\boldsymbol{H}} \boldsymbol{w}_i||^2,$$
$$X = \sum_{c=1}^{C}\sum_{k=1}^{|U_c|} \ln(1+\mu_{c,k}) - \mu_{c,k}, \qquad Y = ||\boldsymbol{\Phi}||^2 \sigma_z^2$$

In order to simplify the objective function in (20), $f_3(\boldsymbol{p})$ can be rewritten as

$$\begin{aligned} f_4(\boldsymbol{p}) = &\sum_{c=1}^{C}\sum_{k=1}^{|U_c|} \bar{A}_{c,k} \mathfrak{Re}\{\bar{B}_{c,k}\}\sqrt{p_{c,k}} \\ &- \sum_{c=1}^{C}\sum_{k=1}^{|U_c|} |v_{c,k}|^2 \Bigg[\sum_{j=1}^{k} |D_{c,k,c}|^2 p_{c,j} \\ &+ \sum_{i\neq c}^{C}\sum_{j=1}^{|U_i|} |D_{c,k,i}|^2 p_{i,j} + E_{c,k}\Bigg] \\ &- \eta\zeta\Bigg(\sum_{i=1}^{C}\sum_{j=1}^{|U_i|} ||\boldsymbol{w}_i||^2 p_{i,j}\Bigg) - Q \\ &- \eta\xi\Bigg(\sum_{i=1}^{C}\sum_{j=1}^{|U_i|} Z_i\, p_{i,j} + Y\Bigg) + X \end{aligned} \tag{21}$$

To solve this optimization problem, we use Lagrange multiplier. The Lagrangian function will be as follows

$$\mathfrak{L}(\boldsymbol{p}, \lambda_1) = f_4(\boldsymbol{p}) - \lambda_1\Big(\sum_{i=1}^{C}\sum_{j=1}^{|U_i|} ||\boldsymbol{w}_i||^2 p_{i,j} - \bar{P}_{max}^{BS}\Big)$$

where λ_1 is the Lagrange multiplier for C_2 constraint, $\lambda_1 \geq 0$. The dual objective function is given by

$$\mathfrak{L}_D(\lambda_1) = \begin{cases} \max\limits_{\boldsymbol{p}} & \mathfrak{L}(\boldsymbol{p}, \lambda_1) \\ \text{s.t} & \text{C2} \end{cases}$$

Then, the dual optimization problem can be written as

$$\lambda_1^* = \min_{\lambda_1 \geq 0} \mathfrak{L}_D(\lambda_1)$$

User optimal power allocation is obtained by differentiating the Lagrange function $\mathfrak{L}(\boldsymbol{p}, \lambda_1)$ with respect to $p_{c,k}$ and λ_1 as

$$\frac{\partial \mathfrak{L}(\boldsymbol{p}, \lambda_1)}{\partial p_{c,k}} = 0, \qquad \frac{\partial \mathfrak{L}(\boldsymbol{p}, \lambda_1)}{\partial \lambda_1} = 0$$

The necessary and sufficient KKT conditions are given as

1) $\frac{\partial \mathfrak{L}(\boldsymbol{p}, \lambda_1)}{\partial p_{c,k}} = 0$
2) $\lambda_1\Big(\sum_{i=1}^{C}\sum_{j=1}^{|U_i|} ||\boldsymbol{w}_i||^2 p_{i,j} - \bar{P}_{max}^{BS}\Big) = 0$
3) $\sum_{i=1}^{C}\sum_{j=1}^{|U_i|} ||\boldsymbol{w}_i||^2 p_{i,j} \leq \bar{P}_{max}^{BS}$
4) $\lambda_1 \;\&\; \boldsymbol{p} \geq 0$

The solution for the optimal user power must satisfy those conditions.

B. OPTIMUM AUXILIARY VARIABLES

Consider fixed user power and LIS phase and gain, the optimal auxiliary variables $(\boldsymbol{\mu}^{opt}, \boldsymbol{v}^{opt})$ are obtained by differentiating $f_2(\boldsymbol{p}, \boldsymbol{\phi}, \boldsymbol{a}, \boldsymbol{\mu}, \boldsymbol{v})$ given in (19) with respect to $\mu_{c,k}$ and $v_{c,k}$ and equate to zero. So, the optimal auxiliary variables are given as follows

$$\mu_{c,k}^{opt} = \frac{\mathfrak{D}_{c,k}}{2}\Big(\mathfrak{D}_{c,k} + \sqrt{\mathfrak{D}_{c,k}^2 + 4}\Big) \tag{22a}$$

$$v_{c,k}^{opt} = \frac{\sqrt{1+\mu_{c,k}} A_{c,k}}{B_{c,k}} \tag{22b}$$

where $\mathfrak{D}_{c,k} = \mathfrak{Re}\{v_{c,k}^* A_{c,k}\}$, $A_{c,k}$ and $B_{c,k}$ are given in (14).

C. OPTIMIZATION OF LIS PHASE AND GAIN

For notation simplicity, let $\boldsymbol{\Phi} = diag(\boldsymbol{\psi})$, $\boldsymbol{\psi}$ is a vector of size $\mathbb{C}^{L\times 1}$ and defined as $\boldsymbol{\psi} = [\psi_1, \psi_2, .., \psi_L]^T = [a_1\phi_1, a_2\phi_2, \ldots, a_L\phi_L]^T$.

In order to get the optimum values for phase shift and gain of LIS elements, problem (19) is formulated as

$$\begin{aligned} \max_{\boldsymbol{\phi},\boldsymbol{a}} \quad & \mathfrak{Re}\{\boldsymbol{\psi}^H \boldsymbol{x}\} - \boldsymbol{\psi}^H \boldsymbol{S} \boldsymbol{\psi} \\ \text{s.t.} \quad & C_3 : \boldsymbol{\psi}^H \boldsymbol{T} \boldsymbol{\psi} \leq \hat{P}_{max}^{RIS} \\ & C_4, C_5 \end{aligned} \tag{23}$$

where

$$\begin{aligned} \boldsymbol{x} = &\sum_{c=1}^{C}\sum_{k=1}^{|U_c|} 2\sqrt{p_{c,k}(1+\mu_{c,k})} v_{c,k}^* diag(\boldsymbol{f}_{c,k}^H)\boldsymbol{\beta}_c \\ &- \sum_{c=1}^{C}\sum_{k=1}^{|U_c|} |v_{c,k}|^2 \Bigg(\sum_{j=1}^{k} 2\alpha_{c,k,c}^* diag(\boldsymbol{f}_{c,k}^H)\boldsymbol{\beta}_c p_{c,j} \\ &+ \sum_{i\neq c}^{C}\sum_{j=1}^{|U_i|} 2\alpha_{c,k,i}^* diag(\boldsymbol{f}_{c,k}^H)\boldsymbol{\beta}_i p_{i,j}\Bigg) \end{aligned}$$

$$\boldsymbol{T} = \sum_{i=1}^{C}\sum_{j=1}^{|U_i|} p_{i,j} diag(\boldsymbol{\beta}_i)(diag(\boldsymbol{\beta}_i))^H + \sigma_z^2 \boldsymbol{I}_L$$

$$\begin{aligned} \boldsymbol{S} = &\sum_{c=1}^{C}\sum_{k=1}^{|U_c|} |v_{c,k}|^2 \sigma_z^2 diag(\boldsymbol{f}_{c,k}^H) diag(\boldsymbol{f}_{c,k}) \\ &+ \sum_{c=1}^{C}\sum_{k=1}^{|U_c|} |v_{c,k}|^2 \Bigg(\sum_{j=1}^{k} p_{c,j} diag(\boldsymbol{\beta}_c^*)\boldsymbol{f}_{c,k}\boldsymbol{f}_{c,k}^H diag(\boldsymbol{\beta}_c) \\ &+ \sum_{i\neq c}^{C}\sum_{j=1}^{|U_i|} p_{i,j} diag(\boldsymbol{\beta}_i^*)\boldsymbol{f}_{c,k}\boldsymbol{f}_{c,k}^H diag(\boldsymbol{\beta}_i)\Bigg) + \eta\xi\boldsymbol{T} \end{aligned}$$

Since, $\boldsymbol{h}_{c,k} = \boldsymbol{g}_{c,k} + \boldsymbol{H}\boldsymbol{\Phi}\boldsymbol{f}_{c,k}$, then, $\boldsymbol{h}_{c,k}^H \boldsymbol{w}_i = \alpha_{c,k,i} + \boldsymbol{f}_{c,k}^H diag(\boldsymbol{\beta}_i)\boldsymbol{\psi}$, where $\alpha_{c,k,i} = \boldsymbol{g}_{c,k}^H \boldsymbol{w}_i$, $\alpha_{c,k,i} \in \mathbb{C}^{1\times 1}$ and $\boldsymbol{\beta}_i = \boldsymbol{H}^H \boldsymbol{w}_i$, $\boldsymbol{\beta}_i \in \mathbb{C}^{L\times 1}$.

Let

$$\hat{P}_{max}^{RIS} = \xi^{-1} \tilde{P}_{max}^{RIS}$$

Note that problem (23) is a standard quadratic constraint quadratic programming (QCQP) problem, so it can be solved using Lagrange multiplier method [34], the optimal solution can be obtained by

$$\boldsymbol{\psi}^{opt} = (\boldsymbol{S} + \lambda_2 \boldsymbol{T})^{-1}\boldsymbol{x} \tag{24}$$

where λ_2 is Lagrange multiplier, which should be chosen such that complementary slackness conditions of power constraint C_3 of the problem (23) is satisfied. The optimal Lagrange multiplier λ_2^{opt} can be obtained by binary search [34]. An initial range for λ_2 is chosen, i.e., $[\lambda_2^{min},\lambda_2^{max}]$. Then, power constraint C_3 of the problem (23) is calculated for λ_2^{min} and λ_2^{max}. If the calculated minimum power is less than the maximum RIS power, then the optimum value is reached. If both the calculated minimum power and maximum power are higher than the maximum RIS power (the constraint is not satisfied), the range is updated and the power is recalculated for the new maximum and minimum λ_2 values. From [31], since η is non decreasing after each iteration, thus convergence of alternating algorithm to local optimal is guaranteed. So, if the updated values of the $\boldsymbol{p}, \boldsymbol{\phi}, \boldsymbol{a}, \boldsymbol{\mu}$ and $\boldsymbol{v}$ are all optimal, a local optimal solution can be obtained by alternating optimizing these parameters until converges of problem (11).

Algorithm 1 Joint Iterative Optimization Algorithm

Input: System parameters and Channels.
Output: Optimized variables $\boldsymbol{p}^{opt}$, $\boldsymbol{\phi}^{opt}$ and $\boldsymbol{a}^{opt}$.
Initialization: $\boldsymbol{p}, \boldsymbol{\phi}, \boldsymbol{a}, \boldsymbol{\mu}, \boldsymbol{v}$ and $t = 1$
1: **while** no convergence of f_1 OR $t \leq T_{max}$ **do**
2: update $\boldsymbol{\mu}^t$ and $\boldsymbol{v}^t$ as in (22)
3: update $\boldsymbol{p}^t$
4: update λ_2 using binary search
5: update $\boldsymbol{\psi}^t$ as in (24)
6: $t = t + 1$
7: **end while**
8: **return** η^{opt} based on $\boldsymbol{p}$ and $\boldsymbol{\psi}$.

The joint iterative optimization algorithm is explained in Algorithm 1. The optimization parameters are first initialized. Then the auxiliary variables are obtained using (22). Afterwards, using the obtained auxiliary variables, the user power allocation is obtained when the necessary and sufficient conditions are satisfied. Then, the phase and gain of LIS elements are derived using (24). The algorithm is repeated from step 2 to 5 until f_1 in (11) is converged or the iteration index t reached the maximum number of iterations T_{max}. Finally, the optimal energy efficiency is obtained after substituting with the optimal parameters.

D. CONVERGENCE AND COMPLEXITY

To illustrate the convergence of Algorithm 1, we show in Fig. 3 the energy efficiency versus the number of iterations for LIS and passive RIS. The figure shows that, for passive RIS, the algorithm converges after about 2 iterations while, for the LIS, it converges after about 5 iterations. The reason behind that, is for the passive RIS, only phase and power are optimized while in LIS, three parameters are optimized, i.e., gain, phase, and power, which take more iterations.

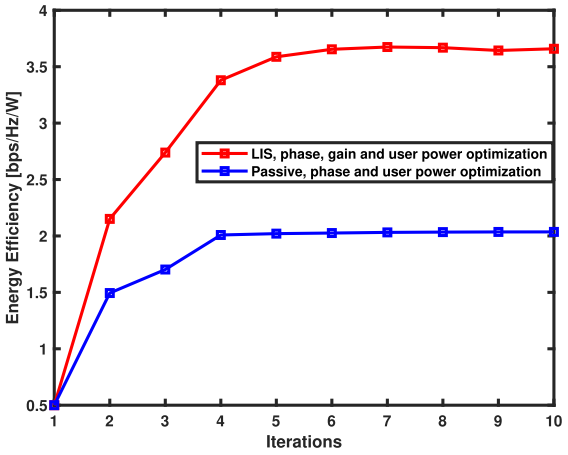

FIGURE 3. Energy efficiency vs. the number iterations (T_{max}) for LIS and passive RIS.

The computational complexity of joint iteration algorithm depends on updating the four variables $\boldsymbol{p}$, $\boldsymbol{\mu}$, $\boldsymbol{v}$ and $\boldsymbol{\psi}$ using (20a),(22a), (22b) and (24) respectively. The complexity of updating $\boldsymbol{p}$ is $\mathcal{O}(M^2N)$ and the complexity of updating $\boldsymbol{\mu}$ is $\mathcal{O}(MN)$. The complexity of updating $\boldsymbol{v}$ is $\mathcal{O}(M^2N + ML)$. For a given tolerance ϖ, the complexity of solving QCQP is $\mathcal{O}(log_2(\frac{1}{\varpi})(\sqrt{L+1}(1+2L)L^3))$. Therefore, the overall complexity of the algorithm is $\mathcal{O}(log_2(\frac{1}{\varpi})T(L^{4.5} + M^2N))$, where T is the number of iterations required for convergence.

In comparison with other computational complexity model, the complexity of LIS phase and gain $\boldsymbol{\psi}$, and the auxiliary variables $\boldsymbol{\mu}$ and $\boldsymbol{v}$ are similar to the complexity computation of Algorithm 1 in [18]. However, the power complexity in our work can be compared with the precoding complexity which is $\mathcal{O}(log_2(\frac{1}{\varpi})T(\sqrt{NM+2}(1 + NM)M^3N^3))$. In addition, the overall complexity of the other model is $\mathcal{O}(log_2(\frac{1}{\varpi})T(M^{4.5}N^{4.5} + L^{4.5}))$. It shows that our algorithm is much lower in complexity than the other model. Assuming M=3, N=4 and L=16, then the overall complexity of our algorithm is 262180 while the other model is 333975.

VII. NUMERICAL RESULTS

In this section, the simulation results are presented. The BS is located at (0,−20,0), RIS is located at (100,5,0) and users are uniform distributed in circle centered at (A,0,0) with radius 5m. The users positions are fixed over the simulations at $A = 200m$, where A represents the horizontal distance from the origin (0,0,0). Users are grouped in clusters, where cluster heads are first chosen then the remaining users are assigned to the nearest cluster head user as described in Section III. From [35], far-field assumption is valid, when the distance between the scattering source and the receiver is larger than Fraunhofer distance d_f. In [34], it is stated that when the

TABLE 1. Simulation parameters.

Description	Values
BS location, (x_1, y_1, z_1)	(0m,-20m,0m)
LIS location, (x_2, y_2, z_2)	(100m,5m,0m)
Number of antennas, N	8
Number of users, M	6
Number of LIS elements, L	256
Number of clusters, C	4
Inverse energy conversion at BS & LIS, (ζ, ξ)	1.1
power dissipated per user, P_u	10mW
static power consumption of phase shifter, P_{PS}	10mW
static power consumption of power amplifier, P_{PA}	10mW
noise variance at user & LIS, (σ, σ_z)	10^{-8}mW
Baseband power at BS, P_B	6dB
maximum power at BS, P_{max}^{BS}	[9dB, 18dB]
maximum power at RIS, P_{max}^{RIS}	9dB

number of elements is less than 361, the far field distance d_f is about 5m. In this work, the distance between BS and RIS is $100m$, the distance between RIS center of users distribution circle is $100m$ and the distance between BS and center of user distribution circle is $200m$. All distances are larger than the far field distance, thus far field is considered.

We consider the BS employs Uniform Linear Array (ULA) with N antennas. Passive RIS is included for comparison purposes. RIS (active or passive) employs Uniform Planar Array (UPA) with $L = L_h L_v$ where L_h is the number of elements along horizontal axis and L_v is the number of elements along vertical axis. The mmWave channel model that is considered in the simulation, has three paths ($P_{c,k} = 3$), one line of sight $\alpha_{c,k}^{(1)}$ and two non-line of sight $\alpha_{c,k}^{(p)}$; $p \in \{2, 3\}$. The paths $\alpha_{c,k}^{(1)}$ and $\alpha_{c,k}^{(p)}$ follow $\mathcal{CN}(0, 10^{-2})$ and $\mathcal{CN}(0, 10^{-3})$ respectively. While $\varphi_{c,k}^{(p)}$ and $\theta_{c,k}^{(p)}$ follow uniform distribution from $[-\pi, \pi]$. The values of the simulation parameters are shown in Table 1.

In this paper, four cases are explored in the simulation:

- Case 1: Optimization of phase, gain of LIS elements and power allocation of users (proposed algorithm).
- Case 2: Optimization of phase, gain of LIS elements only (power is equally divided among users).
- Case 3: Optimization of users powers only (random phase and gain).
- Case 4: Without any optimization (random phase, gain and equally divided power among users).

Energy efficiency comparison of passive RIS and LIS versus (P_{max}^{BS}) is shown in Fig. 4. The results show that as the maximum power dedicated for BS increases, the EE decreases. This is because of the increase of the total power consumption due to the increase of the BS power. Moreover, it can be observed that the EE obtained by jointly optimizing user powers, phase and gain of LIS elements has superior performance when compared to the other optimization cases. The reason behind that is that by jointly optimizing phase, gain with power leads to minimize the power consumption which reflects back on the EE and shows better performance

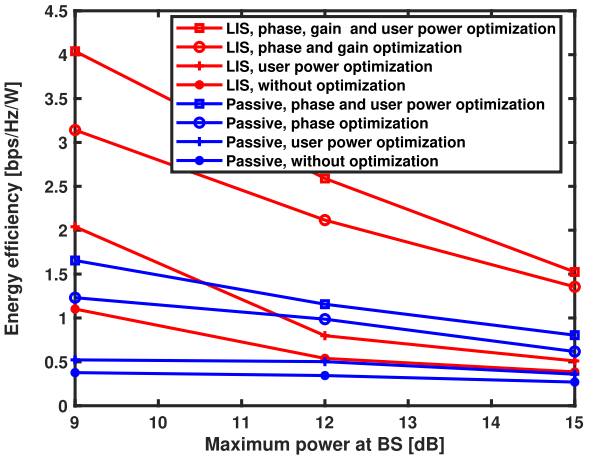

FIGURE 4. Energy efficiency vs. maximum power at BS (P_{max}^{BS}) for LIS and passive RIS.

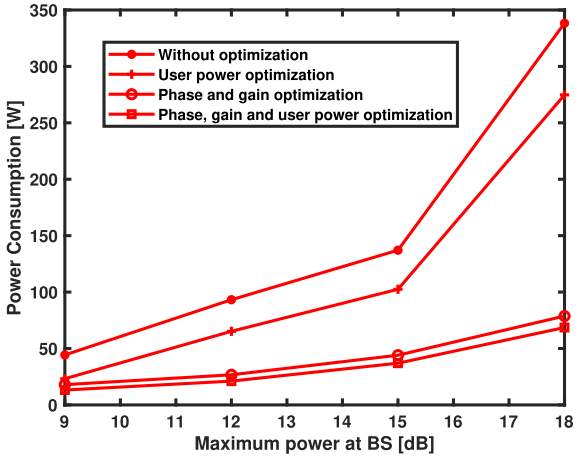

FIGURE 5. Power consumption vs. maximum power at BS (P_{max}^{BS}) for different cases of LIS.

than optimizing only the phase and gain. LIS in any case outperforms passive RIS due to optimization of the gain. Similar to LIS, phase of passive RIS elements and user power optimization (case 1) has superior EE performance when compared to the other optimization cases of passive RIS. There is a slight performance enhancement when user power optimization (case 3) compared to case 4 (without optimization). This is due to optimize allocating power to the users is better than just using fixed power allocation.

Fig. 5 illustrates the total power consumption of LIS versus the maximum power at BS. It is shown that, the power consumption increases with increasing the maximum power. The lowest power consumption is achieved by joint optimization of phase, gain and user power which confirms the results of Fig.4. This indicates how optimization enhances the overall performance. This is also shown in the figure: case 4 has the highest power consumption.

A comparison of the sum rate for LIS and passive RIS is shown in Fig. 6. As previously mentioned, as the maximum power dedicated for BS increases, the sum rate increases.

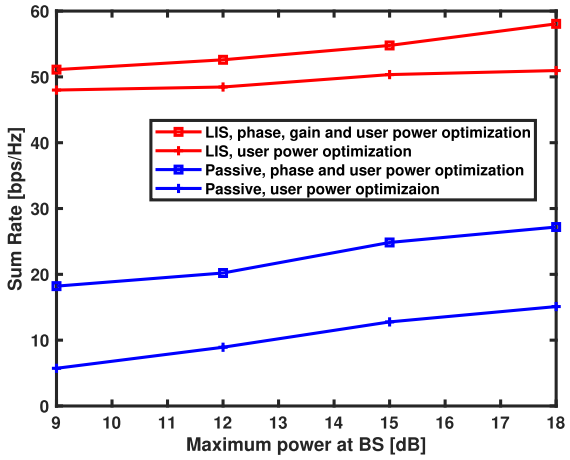

FIGURE 6. Sum rate vs. maximum power at BS (P_{max}^{BS}) for LIS and passive RIS.

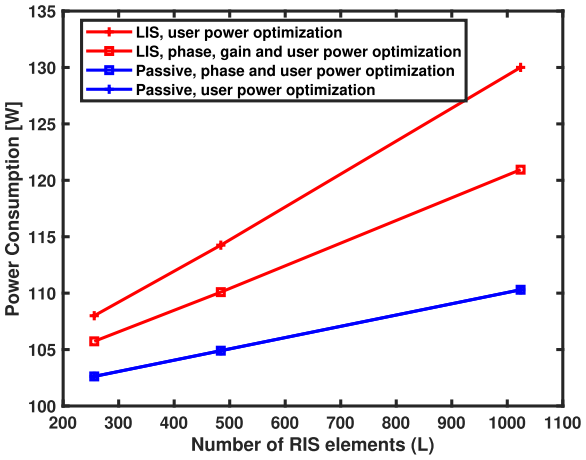

FIGURE 8. Power consumption vs. number of RIS elements (L) for LIS and passive RIS.

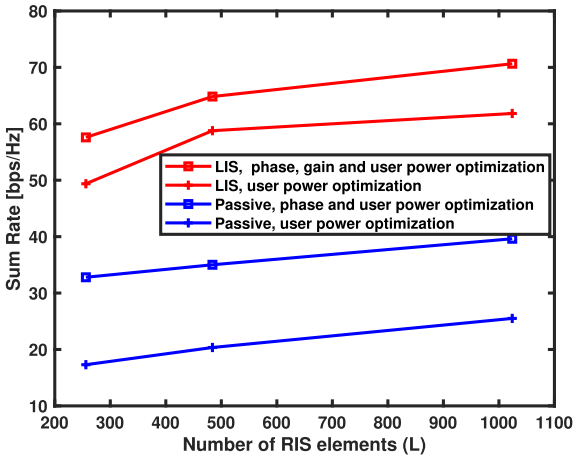

FIGURE 7. Sum rate vs. number of RIS elements (L) for LIS and passive RIS.

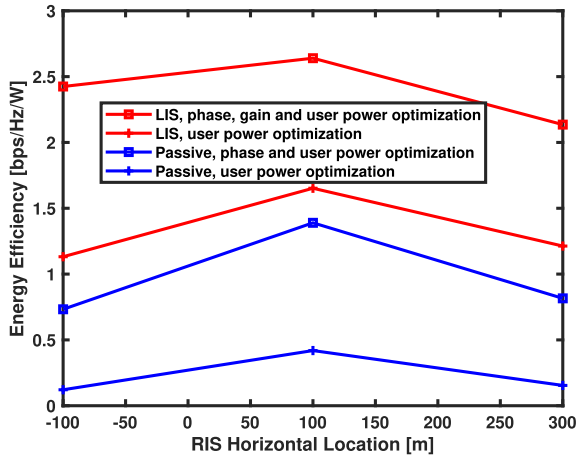

FIGURE 9. Energy efficiency vs. RIS horizontal location for LIS and passive RIS.

For passive RIS, phase and user power optimization (case 1) achieves higher sum rate than user power optimization only (case 3). The gain in sum rate for passive RIS in case 1 is 219% relative to case 3. For optimization of the user powers, LIS gain and phase (case 1) outperforms the optimization of user powers only (case 3). The figure shows that the sum rate for LIS is higher than the sum rate for passive RIS for both cases. The figure also shows that, there is an increases in the LIS (case 1) relative to passive RIS (case 1) by 183% at $P_{max}^{BS} = 9$dB.

Sum rate versus the number of RIS elements (L) for active and passive RIS is illustrated in Fig. 7. As the number of RIS elements increases, the sum rate increases. This is because as L increases, the reflected links to the users are increased and then the received signal is enhanced. In addition, sum rate of LIS outperforms passive RIS. This is due to the effect of the gain of the power amplifier of the LIS. Moreover as mentioned before, the optimization of RIS elements and user power give better sum rate when compared to the optimization of user power only for active/passive RIS.

The tradeoff between active and passive RIS in terms of power consumption is shown in Fig. 8. The figure shows that LIS suffers from high power consumption relative to passive RIS. This results from power dissipated in power amplifier in LIS. It is clear that as the number of RIS elements increases, the total power consumption increases. Since in passive RIS the phase norm equals to one, thus the passive RIS power consumption for phase and user power optimization (case 1) is the same as the power consumption for user power optimization (case 3) indicated by two overlapping curves. However in LIS, a random gain for user power optimization (case 3) is considered, which leads to higher power consumption compared to case 1.

Both Fig. 9 and Fig. 10 illustrate the energy efficiency and sum rate versus the RIS horizontal location for LIS/passive RIS respectively. It is shown that when RIS is located at 100 m in the x-axis, the curves achieve high EE and sum rate relative to the other locations, which indicates that in this scenario, the optimum location for the LIS is at 100m (mid way between the BS and the center of the users distribution circle).

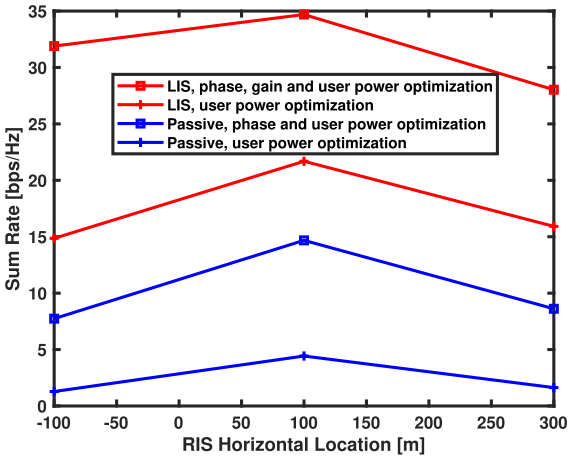

FIGURE 10. Sum rate vs. RIS horizontal location for LIS and passive RIS.

VIII. CONCLUSION

In this paper, we propose the joint optimization of LIS elements (phase, gain) and NOMA user powers. The performance of LIS and passive RIS is investigated in mmWave multi-user MIMO-NOMA system. Results show that LIS outperforms passive RIS in spectral efficiency. In contrast, the total power consumption in LIS is high compared to passive RIS. However, LIS still enjoys higher energy efficiency than passive RIS. Thus, the joint optimization of LIS elements have a significant effect on the overall system performance. For future work, a hybrid LIS may be considered, in which some elements are passive and others are active. This will decrease the total power consumption of the system. In future work, we will also consider user fairness by updating the minimum rate constraint for each user.

REFERENCES

[1] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, "Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 870–913, 2nd Quart., 2018.

[2] M. Xiao, S. Mumtaz, Y. Huang, L. Dai, Y. Li, M. Matthaiou, G. K. Karagiannidis, E. Björnson, K. Yang, C.-L. I, and A. Ghosh, "Millimeter wave communications for future mobile networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1909–1935, Sep. 2017.

[3] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath Jr., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[4] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, "A survey of non-orthogonal multiple access for 5G," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2294–2323, 3rd Quart., 2018.

[5] 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.

[6] T. Sharma, A. Chehri, and P. Fortier, "Reconfigurable intelligent surfaces for 5G and beyond wireless communications: A comprehensive survey," *Energies*, vol. 14, no. 24, p. 8219, Dec. 2021.

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

[8] Q. Gu, D. Wu, X. Su, J. Jin, Y. Yuan, and J. Wang, "Performance comparison between reconfigurable intelligent surface and relays: Theoretical methods and a perspective from operator," 2021, *arXiv:2101.12091*.

[9] M. Jian, G. C. Alexandropoulos, E. Basar, C. Huang, R. Liu, Y. Liu, and C. Yuen, "Reconfigurable intelligent surfaces for wireless communications: Overview of hardware designs, channel models, and estimation techniques," *Intell. Converged Netw.*, vol. 3, no. 1, pp. 1–32, Mar. 2022.

[10] M. A. ElMossallamy, H. Zhang, L. Song, K. G. Seddik, Z. Han, and G. Y. Li, "Reconfigurable intelligent surfaces for wireless communications: Principles, challenges, and opportunities," *IEEE Trans. Cognit. Commun. Netw.*, vol. 6, no. 3, pp. 990–1002, Sep. 2020.

[11] Y. Cao, T. Lv, and W. Ni, "Intelligent reflecting surface aided multi-user mmWave communications for coverage enhancement," in *Proc. IEEE 31st Annu. Int. Symp. Pers., Indoor Mobile Radio Commun.*, Aug. 2020, pp. 1–6.

[12] L. Zhang, Q. Wang, and H. Wang, "Multiple intelligent reflecting surface aided multi-user weighted sum-rate maximization using manifold optimization," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Jul. 2021, pp. 364–369.

[13] J. Li and J. Liu, "Sum rate maximization via reconfigurable intelligent surface in UAV communication: Phase shift and trajectory optimization," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, China, Aug. 2020, pp. 124–129.

[14] B. Guo, C. Sun, and M. Tao, "Two-way passive beamforming design for RIS-aided FDD communication systems," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2021, pp. 1–6.

[15] T. Jiang and W. Yu, "Interference nulling using reconfigurable intelligent surface," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 5, pp. 1392–1406, May 2022.

[16] P. Wang, J. Fang, X. Yuan, Z. Chen, and H. Li, "Intelligent reflecting surface-assisted millimeter wave communications: Joint active and passive precoding design," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 14960–14973, Dec. 2020.

[17] K. Liu, Z. Zhang, L. Dai, S. Xu, and F. Yang, "Active reconfigurable intelligent surface: Fully-connected or sub-connected?" *IEEE Commun. Lett.*, vol. 26, no. 1, pp. 167–171, Jan. 2022.

[18] Z. Zhang, L. Dai, X. Chen, C. Liu, F. Yang, R. Schober, and H. V. Poor, "Active RIS vs. passive RIS: Which will prevail in 6G?" 2021, *arXiv:2103.15154*.

[19] 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.

[20] Y. Lu, M. Hao, and R. Mackenzie, "Reconfigurable intelligent surface based hybrid precoding for THz communications," *Intell. Converged Netw.*, vol. 3, no. 1, pp. 103–118, Mar. 2022.

[21] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-lin, and Z. Wang, "Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sep. 2015.

[22] A. S. D. Sena, D. Carrillo, F. Fang, P. H. J. Nardelli, D. B. D. Costa, U. S. Dias, Z. Ding, C. B. Papadias, and W. Saad, "What role do intelligent reflecting surfaces play in multi-antenna non-orthogonal multiple access?" *IEEE Wireless Commun.*, vol. 27, no. 5, pp. 24–31, Oct. 2020.

[23] Y. Liu, X. Mu, X. Liu, M. Di Renzo, Z. Ding, and R. Schober, "Reconfigurable intelligent surface-aided multi-user networks: Interplay between NOMA and RIS," *IEEE Wireless Commun.*, vol. 29, no. 2, pp. 169–176, Apr. 2022.

[24] P. Liu, Y. Li, W. Cheng, X. Gao, and X. Huang, "Intelligent reflecting surface aided NOMA for millimeter-wave massive MIMO with lens antenna array," *IEEE Trans. Veh. Technol.*, vol. 70, no. 5, pp. 4419–4434, May 2021.

[25] A. Ihsan, W. Chen, M. Asif, W. Ullah Khan, and J. Li, "Energy-efficient IRS-aided NOMA beamforming for 6G wireless communications," 2022, *arXiv:2203.16099*.

[26] A. Khaleel and E. Basar, "A novel NOMA solution with RIS partitioning," *IEEE J. Sel. Topics Signal Process.*, vol. 16, no. 1, pp. 70–81, Jan. 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. 14088–14092, Nov. 2020.

[28] L. N. Ribeiro, S. Schwarz, and M. Haardt, "Low-complexity zero-forcing precoding for XL-MIMO transmissions," in *Proc. 29th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2021, pp. 1621–1625.

[29] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059–3075, Jun. 2015.
[30] J.-F. Bousquet, S. Magierowski, and G. G. Messier, "A 4-GHz active scatterer in 130-nm CMOS for phase sweep Amplify-and-Forward," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 59, no. 3, pp. 529–540, Mar. 2012.
[31] K. Shen and W. Yu, "Fractional programming for communication systems—Part I: Power control and beamforming," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2616–2630, May 2018.
[32] K. Shen and W. Yu, "Fractional programming for communication systems—Part II: Uplink scheduling via matching," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2631–2644, May 2018.
[33] Y. Xu and W. Yin, "A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion," *SIAM J. Imag. Sci.*, vol. 6, no. 3, pp. 1758–1789, Jan. 2013.
[34] S. Boyd, *Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers* (Foundations and Trends in Machine Learning), vol. 3:1. Boston, MA, USA: Now, 2010.
[35] G. C. Alexandropoulos, V. Jamali, R. Schober, and H. V. Poor, "Near-field hierarchical beam management for RIS-enabled millimeter wave multi-antenna systems," in *Proc. IEEE 12th Sensor Array Multichannel Signal Process. Workshop (SAM)*, Jun. 2022, pp. 460–464.

NOURAN ARAFAT was born in Cairo. She received the B.S. and M.S. degrees in information engineering and technology (IET), communications major from German University in Cairo, Egypt, in 2013 and 2015, respectively, where she is currently pursuing the Ph.D. degree in information engineering and technology. Her research interests include wireless communications, mmWave massive MIMO, NOMA, and deep learning-based wireless communications.

ENGY ALY MAHER (Senior Member, IEEE) was born in Cairo, Egypt, in 1990. She received the B.S. degree (Hons.) in information engineering and technology (IET), communications major from German University in Cairo (GUC), Cairo, in 2012, the M.S. degree, in 2013, and the Ph.D. degree from GUC, in 2019. From 2012 to 2019, she was a Teaching and Research Assistant with the Electrical Engineering Department, GUC, where she has been an Assistant Professor and a Researcher, since 2019. Her research interests include mobile and wireless communications, resource allocation, power control, radio resource management, spectral efficiency, and signal processing. Her work is recognized by several publications at both national and international conferences and journals in this field.

AHMED EL-MAHDY (Senior Member, IEEE) was born in Port Said, Egypt, in 1963. He received the B.S. and M.S. degrees in electrical engineering, in 1986 and 1992, respectively, and the Ph.D. degree in electrical engineering from the Catholic University of America, Washington, DC, USA, in 1999. From 1986 to 1989, he was an Electronic Engineer. From 1989 to 1992, he was a Researcher and an Instructor with the Electrical Engineering Department, MTC, Cairo, Egypt, where he has been a Lecturer and a Researcher, since 1999. He was promoted to an Associate Professor, in 2005. He is currently a Professor and the Dean of the Faculty of Information Engineering & Technology, German University in Cairo, Egypt. His research interests include wireless communications, equalization, detection and estimation, and adaptive signal processing. He received the National Encouragement Prize in Engineering Sciences for Wireless Communications and Signal Processing Researches, in June 2008.

FALKO DRESSLER (Fellow, IEEE) received the M.Sc. and Ph.D. degrees from the Department of Computer Science, University of Erlangen, in 1998 and 2003, respectively. He is currently a Full Professor and the Chair of Telecommunication Networks with the School of Electrical Engineering and Computer Science, TU Berlin. He has authored the textbooks *Self-Organization in Sensor and Actor Networks* (Wiley & Sons) and *Vehicular Networking* (Cambridge University Press). His research interests include adaptive wireless networking (sub-6GHz, mmWave, visible light, and molecular communication) and wireless-based sensing with applications in ad-hoc and sensor networks, the Internet of Things, and cyber-physical systems.

He has been the Associate Editor-in-Chief of IEEE Transactions on Mobile Computing and *Computer Communications* (Elsevier) and an Editor of journals, such as IEEE/ACM Transactions on Networking, IEEE Transactions on Network Science and Engineering, *Ad Hoc Networks* (Elsevier), and *Nano Communication Networks* (Elsevier). He has been chairing conferences, such as IEEE INFOCOM, ACM MobiSys, ACM MobiHoc, IEEE VNC, and IEEE GLOBECOM. He has been an IEEE Distinguished Lecturer and an ACM Distinguished Speaker. He is an ACM Distinguished Member. He is a member of the German National Academy of Science and Engineering (acatech). He has been serving on the IEEE COMSOC Conference Council and the ACM SIGMOBILE Executive Committee.