View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by UCL Discovery

Green Communication for NOMA-Based Cloud Radio Access Network

Wanming Hao, *Student Member, IEEE,* Zheng Chu, *Member, IEEE,* Fuhui Zhou, *Member, IEEE,* Shouyi Yang, and Gangcan Sun

***Abstract*—The wireless devices will rapidly growth with the application of the Internet of Thing (IoT), which results in the huge power consumption. To realize the green communication, in this paper, we study the energy efficiency (EE) problem in a non-orthogonal multiple access (NOMA)-based CRAN, where microwave and millimeter wave (mmWave) are applied in fronthaul and access links, respectively. Based on this, we formulate the power optimization problem to maximize the EE of the system subject to the fronthaul capacity and transmit power constraints. Due to the formulated problem is nonconvex, we first transform the fractional objective function into the subtractive form. After that, an algorithm containing outer and inner loops is proposed to deal with this non-convex problem. In particular, in the outer loop, the ℓ_1-norm technique is adopted to transform the nonconvex fronthaul capacity constraint into the convex one; while in the inner loop, the weighted minimum mean square error (WMMSE) approach is applied to solve the formulated problem. Simulation results demonstrate that the proposed NOMA scheme can obtain higher EE as well as throughput in comparison with the conventional orthogonal multiple access (OMA) scheme.**

***Index Terms*—IoT, green communication, EE, CRAN, mmWave, NOMA**

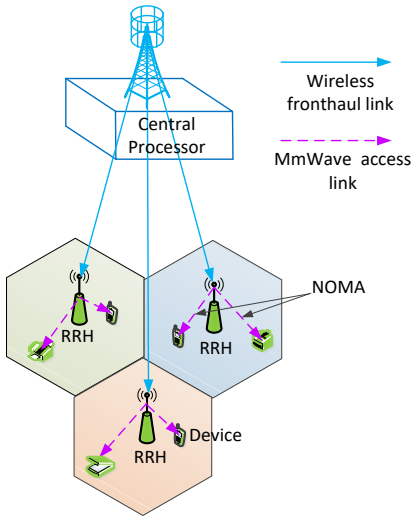

Fig. 1: The NOMA-based cloud radio access network.

I. Introduction

With the application and development of the Internet of Things (IoT), next generation wireless communication systems are expected to support billions of wireless devices, including smartphones, cars, electronic households and so on [1]. The huge number of wireless devices leads to the growing energy consumption of wireless communications, which brings the increase of the greenhouse. Therefore, green communication (i.e., energy-efficient communication) has became a significant standard for future IoT networks, which means that the low energy consumption and high data throughput should be simultaneously satisfied [2], [3].

This work was supported by the National Natural Science Foundation of China under Grant U1604159 and the Special Project for Inter-government Collaboration of State Key Research and Development Program under Grant 2016YFE0118400.

W. Hao, S. Yang and Gangcan Sun are with the School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China, and W. Hao also with the Graduate School of Information Science and Electrical Engineering, Kyushu University, Fukuoka 819-0395, Japan. (E-mails: wmhao@hotmail.com, iesyyang@zzu.edu.cn, iegcsun@zzu.edu.cn)

Z. Chu is with the 5G Innovation Centre, Institute of Communication Systems, University of Surrey, Guildford GU2 7XH, U.K. (Email: zheng.chu@surrey.ac.uk)

F. Zhou is with the Department of Electrical and Computer Engineering, Utah State University, Logan, UT, USA. F. Zhou is also with the School of Information Engineering and Post-Doctoral Research Station of Environmental Science and Engineering, Nanchang University, Nanchang 330031, China. (Email: zhoufuhui1989@163.com)

To realize the requirement of the massive wireless devices connectivity in IoT networks, non-orthogonal multiple access (NOMA) technique can be as an effective candidate because it allows multiple devices to connect to the network with the same time frequency resource [4]. Furthermore, the NOMA technique has recently been included into the 3GPP long term evolution advanced (LTE-A) standard due to its enormous potential in improving the spectrum efficiency (SE) of the system [5], [6]. On the other hand, the cloud radio access network (CRAN), where the conventional base stations (BSs) with large coverage area are replaced by remote radio heads (RRHs) with low power and small coverage, has been as a promising scheme to obtain the high SE and energy efficiency (EE) for future wireless networks [7]. In CRAN, the complicated signal processing and resource allocation are conducted by cloud with a central processor (CP), and RRHs only take charge of the radio frequency processing. Based on this, the low-complexity RRHs and centralized resource allocation simplify the system management and enhance the system performance [8]. Therefore, the combination of the NOMA and CRAN to form a NOMA-based CRAN (as shown in Fig. 1) will be a perfect candidate for future IoT networks.

A. Related Works

Recently, CRAN has been obtained the great attention by researchers and some related study results have been published.

In [9], the authors aim to minimize the total transmit power of the RRHs by optimizing the set of the active RRHs, the precoding and transmit power while maintaining the fronthaul capacity and devices' quality of service (QoS) constraints. Then, two algorithms, the pricing-based and linear-relaxed algorithms, are proposed to solve the formulated problem. The authors in [10] study the joint decompression and decoding for an uplink CRAN and propose an iterative algorithm to maximize the achievable uplink sum rate. In [11], the system utility is first defined and then the authors focus on joint optimization of device grouping and transmit beamforming to maximize the system utility subject to the devices' QoS and the RRHs' power constraints. To avoid the high computational complexity, a low-complexity two-stage iterative algorithm is proposed. In [12], the authors consider the delay problem in uplink CRAN and a low-complexity delay-aware fronthaul allocation algorithm is proposed for minimizing the system delay. A downlink CRAN utility maximization problem is formulated in [13] by optimizing the device scheduling, BS clustering and beamforming design. Based on the above, the authors propose a two-stage iterative algorithm to solve the original optimization problem. Meanwhile, some advanced virtual resource (e.g., BS and antennas) sharing schemes are proposed to improve the performance of the CRAN [14], [15]. However, the above works all consider the wired fronthaul links (e.g., fiber). For the deployment of the ultra-dense RRHs, it is impractical and of high cost to have wired connections between CP and all RRHs.

In fact, there have been some works considering the wireless fronthaul links. In [8], the authors apply the millimeter-wave (mmWave) wireless fronthaul links and orthogonal frequency division multiple access (OFMDA)-based access links. Then, a joint power and subcarrier allocation algorithm is proposed to maximize the weighted sum rate of the devices. In [16], the wired and wireless fronthaul are simultaneously adopted. Accordingly, the authors define an economical SE (ESE) metric to jointly consider the impact of wired/wireless fronthaul cost and traditional EE, and an outer- and inner-based iterative algorithm is proposed for maximizing ESE. In [17], the EE maximization problem is considered in a heterogeneous CRAN with constraint of overall capacity of wireless fronthaul, and a convex relaxation-based power allocation algorithm is proposed. The authors in [18] study a joint design framework of fronthaul and access links to maximize the sum rate of the CRAN. Meanwhile, the multi-device beamforming scheme is adopted for exploiting the spatial diversity. Then, a difference of convex (DOC) and weighted minimum-mean-squared-error (WMMSE)-based algorithm is proposed. Although the above works consider the wireless fronthaul links, the fronthaul capacity are usually assumed as fixed (e.g., [16], [17]). In addition, the transmit power of the CP is also considered a constant in [8]. [18] only considers the system throughput and does not involve EE.

It is clear that [8]-[18] only focus on the resource allocation in CRAN. As a significant technique candidate for the future IoT networks, NOMA-based CRAN has been not obtained the enough attention. In [19], the authors analyze the outage probability of the NOMA-based CRAN. Similarly, the authors in [20] derive a simple expression in terms of outage probability for both nearby devices and the cell-edge devices. To filling this research gap, in this paper, we study the EE problem in a NOMA-based CRAN with wireless fronthaul links.

B. Main Contributions

To realize the green communication for future wireless communication, we study the EE problem in a NOMA-based CRAN, where the CP is equipped with a large number of antennas and the microwave-based wireless fronthaul link is considered. Due to the small coverage area of the RRHs, the mmWave-based access link is adopted. Meanwhile, all RRHs work in full duplex (FD) pattern, i.e., they can simultaneously receive data from the CP through microwave and transmit data to devices through mmWave. Meanwhile, all devices access the RRHs with NOMA for improving the SE. Based on the above, we formulate a power optimization problem (including the CP's and RRHs' transmit power) to maximize the EE of the system. The main contributions are summarized as follows:

- We construct a NOMA-based CRAN system, where the microwave and mmWave are used for frothaul and access links, respectively. Then, we formulate an EE maximization problem by optimizing the CP's and RRHs' transmit power while satisfying the devices' QoS and the constraints of the CP's and RRHs' transmit power.

- By exploiting fractional programming, we first transform the original EE-based fractional objective function into the subtractive form by bringing the parameter. However, the parameter-based objective function is still nonconcave. Furthermore, the non-convex fronthaul capacity constraint also leads that the solved problem is a conconvex optimization problem. Based on this, we propose a two-loop iterative algorithm to deal with it. Specifically, a reweighted ℓ_1-norm technique-based outer iterative is applied to transform the non-convex fronthaul capacity constraint into a convex one. After that, the WMMSE approach-based inner iterative is designed to transform the solved problem into a convex optimization problem, which is solved by standard convex optimization techniques.

- We show numerically the fast convergence and effectiveness of the proposed algorithm. Meanwhile, compared with the conventional orthogonal multiple access (OMA) scheme, our proposed NOMA scheme can obtain a higher EE as well as throughput of the system.

The remainder of this paper is organized as follows. In Section II, the NOMA-based CRAN system model is described and EE maximization problem is formulated. In Section III, the iterative algorithm is proposed for solving the formulated problem. Numerical results are presented in Section IV and the paper is concluded in Section V.

The notations of this paper are as follows: $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ denote the conjugate, transpose and Hermitian transpose, respectively, $\|\cdot\|_n$ means the n norm, $\mathbb{E}\{\cdot\}$ represents the expectation operator, $\mathrm{Re}(\cdot)$ means the real number operation,

$[\cdot]^+$ denotes the $\max\{0,\cdot\}$. A summary of key notations is presented in Table I.

II. System Model and Problem Formulation

A. System Model

The general architecture of the NOMA-based CRAN is showed in Fig. 1, where the RRHs receive the devices' data from the CP via microwave communication. Meanwhile, the RRHs forward the received data to devices by mmWave communication. In this paper, we use the FD technology for fronthaul links, where the FD communication hardware is equipped for each RRH. As the mmWave access and microwave fronthaul transmissions are over different frequency bands, there are no interference when the RRHs simultaneously receive data from the CP and transmit them to devices. We assume that there are N single-antenna RRHs, and K single-antenna devices associate to each RRH with the NOMA while the CP is equipped with M antennas ($M \gg N$). We assume that the perfect channel state information (CSI) are available at CP for central signal precoessing and resource allocation.

1) The Microwave Fronthaul Link: The received signal at the nth RRH is given by

$$y_n^{\mathrm{FH}} = \underbrace{\mathbf{h}_n \sqrt{P_n^{\mathrm{FH}}}\mathbf{v}_n x_n}_{\text{Desired signal}} + \underbrace{\mathbf{h}_n \sum_{l\neq n}^{N} \sqrt{P_l^{\mathrm{FH}}}\mathbf{v}_l x_l}_{\text{Multi-RRH interference}} + \underbrace{z_n}_{\text{Noise}}, \tag{1}$$

where $\mathbf{h}_n \in \mathbb{C}^{1\times M}$ denotes the downlink fronthaul channel from the CP to the nth RRH, while P_n^{FH}, $\mathbf{v}_n \in \mathbb{C}^{M\times 1}$ and x_n represent the transmit power, precoding vector and signal for the nth RRH, respectively. z_n is an independent and identically distributed (i.i.d.) additive while Gaussian noise (AWGN) defined as $\mathcal{CN}(0, N_0)$. To cancel the multi-RRH interference, the classical zero forcing (ZF) precoding is applied and thus we have $\mathbf{V} = \mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1}$, where $\mathbf{H} = [\mathbf{h}_1^T, ..., \mathbf{h}_N^T]^T$. To this end, the precoding of the nth RRH can be expressed as $\mathbf{v}_n = \mathbf{V}_n/\|\mathbf{V}_n\|$, where $\mathbf{V}_n$ is the nth column of $\mathbf{V}$. Accordingly, the achievable throughput of the nth RRH can be expressed as

$$R_n^{\mathrm{FH}}(P_n^{\mathrm{FH}}) = W_{\mathrm{FH}}\log_2\left(1 + \frac{\|\mathbf{h}_n\mathbf{v}_n\|^2 P_n^{\mathrm{FH}}}{W_{\mathrm{FH}}N_0}\right), \tag{2}$$

where W_{FH} is the overall bandwidth in microwave communications.

2) The MmWave Access Link: We denote g_{njk} as the channel coefficient from the nth RRH to the kth device served by the jth RRH. Without the generality, the channels are sorted as $|g_{nn1}|^2 \leq \cdots \leq |g_{nnK}|^2$ according to their channel quality. According to the NOMA protocol, the received signal of the kth device served by the nth RRH can be written as

$$y_{nk}^{\mathrm{AC}} = \underbrace{g_{nnk}\sum_{i=1}^{K}\sqrt{P_{ni}^{\mathrm{AC}}}s_{ni}}_{\text{Desired signal}} + \underbrace{\sum_{j\neq n}^{N} g_{jnk}\sum_{i=1}^{K}\sqrt{P_{ji}^{\mathrm{AC}}}s_{ji}}_{\text{Inter-RRH interference}} + \underbrace{z_{nk}}_{\text{Noise}}, \tag{3}$$

where s_{nk} denotes the transmit signal for the kth device served by the nth RRH, and z_{nk} is an i.i.d. AWGN defined as $\mathcal{CN}(0, N_0)$. After that, the successive interference cancellation technique will be carried out at the devices. As a result, the kth device will detect and decode the ith ($i > m$) device's data and remove it from its observation; while other devices' data will be treated as noise. Accordingly, the received signal can be rewritten as

$$y_{nk}^{\mathrm{AC}} = \underbrace{g_{nnk}\sum_{i=1}^{k}\sqrt{P_{ni}^{\mathrm{AC}}}s_{ni}}_{\text{Desired signal}} + \underbrace{\sum_{j\neq n}^{N} g_{jnk}\sum_{i=1}^{K}\sqrt{P_{ji}^{\mathrm{AC}}}s_{ji}}_{\text{Inter-RRH interference}} + \underbrace{z_{nk}}_{\text{Noise}}, \tag{4}$$

and the achievable throughput can be expressed as

$$R_{nk}^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}}) = W_{\mathrm{AC}}\log_2\left(1 + \frac{|g_{nnk}|^2 P_{nk}^{\mathrm{AC}}}{\pi_{nk}}\right), \tag{5}$$

where $\pi_{nk} = |g_{nnk}|^2\sum_{i=1}^{k-1}P_{ni}^{\mathrm{AC}} + \sum_{j\neq n}^{N}|g_{jnk}|^2\sum_{i=1}^{K}P_{ji}^{\mathrm{AC}} + W_{\mathrm{AC}}N_0$, $\mathcal{P}_{\mathrm{AC}}$ denotes the power allocation policy of all RRHs (i.e. $\{P_{nk}^{\mathrm{AC}}\}_{N\times K}$), and W_{AC} is the overall bandwidth in mmWave communications.

B. Problem Formulation

In general, the power consumption consists of transmit power and circuit power consumption. To this end, the power consumption at the CP can be modeled as

$$P^{\mathrm{FH}}(\mathcal{P}_{\mathrm{FH}}) = \xi\sum_{n=1}^{N}P_n^{\mathrm{FH}} + P_c^{\mathrm{FH}}, \tag{6}$$

where $\mathcal{P}_{\mathrm{FH}}$ denotes the power allocation policy of the CP (i.e. $\{P_{nk}^{\mathrm{AC}}\}_{N\times 1}$), ξ is the reciprocal of drain efficiency of power amplifier, and P_c^{FH} denotes the circuit power consumption. Similarly, the power consumption at RRHs can expressed as

$$P^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}}) = \xi\sum_{n=1}^{N}\sum_{k=1}^{K}P_{nk}^{\mathrm{AC}} + \sum_{n=1}^{N}P_{c,n}^{\mathrm{AC}}, \tag{7}$$

where $P_{c,n}^{\mathrm{AC}}$ denotes the circuit power consumption of the nth RRH. To this end, we define the EE as

$$\eta_{\mathrm{EE}}(\mathcal{P}_{\mathrm{FH}}, \mathcal{P}_{\mathrm{AC}}) = \frac{\sum_{n=1}^{N}\sum_{k=1}^{K}R_{nk}^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}})}{P^{\mathrm{FH}}(\mathcal{P}_{\mathrm{FH}}) + P^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}})}[\text{bps/Joule}], \tag{8}$$

TABLE I: Summary of Key Notations.

Notations	Descriptions
N	Number of RRHs
K	Number of devices per RRH
M	Number of CP antennas
W_{FH}	Bandwidth of the microwave
W_{AC}	Bandwidth of the mmWave
P_n^{FH}	Transmit power for the nth RRH
$P_{\max}^{\mathrm{FH}}$	Maximal transmit power for the nth RRH
P_{nk}^{AC}	Transmit power for the kth device served by the nth RRH
$P_{n,\max}^{\mathrm{AC}}$	Maximal transmit power for the nth RRH
P_c^{FH}	Circuit power consumption for the CP
$P_{c,n}^{\mathrm{AC}}$	Circuit power consumption for the nth RRH
$R_{nk,\min}^{\mathrm{AC}}$	The QoS for the kth device served by the nth RRH
$\mathcal{P}^{\mathrm{FH}}$	Power allocation policy of the CP
$\mathcal{P}^{\mathrm{AC}}$	Power allocation policy of the RRHs
ξ	Drain inefficiency of the power amplifier

and formulate the EE maximization problem as follows

$$\max_{\{P_n^{\mathrm{FH}}, P_{nk}^{\mathrm{AC}}\}} \quad \eta_{\mathrm{EE}}(\mathcal{P}_{\mathrm{FH}}, \mathcal{P}_{\mathrm{AC}}), \tag{9a}$$

$$\text{s.t. } R_{nk}^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}}) \geq R_{nk,\min}^{\mathrm{AC}}, \forall n, k \tag{9b}$$

$$\sum_{n=1}^{N} P_n^{\mathrm{FH}} \leq P_{\max}^{\mathrm{FH}}, \tag{9c}$$

$$\sum_{k=1}^{K} P_{nk}^{\mathrm{AC}} \leq P_{n,\max}^{\mathrm{AC}}, n \in \{1, \cdots, N\}, \tag{9d}$$

$$\sum_{k=1}^{K} R_{nk}^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}}) \leq R_n^{\mathrm{FH}}(\mathcal{P}_{\mathrm{FH}}), n \in \{1, \cdots, N\}, \tag{9e}$$

$$P_n^{\mathrm{FH}} \geq 0, P_{nk}^{\mathrm{AC}} \geq 0, \forall n, k, \tag{9f}$$

where (9b) denotes the device's minimal throughput requirement, while (9c) and (9d) denote the power constraints for the CP and each RRH, respectively. (9e) denotes the fronthaul capacity constraint, which means that the fronthaul capacity of each RRH must be no less than the capacity it provides for K devices. Due to the fractional objective function (9b), inter-RRH interference and the non-convex constraints (9e), (9) is a non-convex optimization problem.

III. The Solution of The Optimization Problem

We solve the original EE optimization problem by three steps. At the first step, we equivalently transform the original fractional objective function into a subtractive form. Next, the non-convex fronthaul capacity constraint is approximated as convex one by ℓ_1-norm technique. Finally, we apply the WMMSE approach to solve the formulated problem.

A. The Transformation of Objective Function

We assume that the maximum EE of the problem (9) is $\eta^\star$, namely

$$\begin{aligned} \eta^\star &= \frac{\sum_{n=1}^{N}\sum_{k=1}^{K} R_{nk}^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}}^\star)}{P^{\mathrm{FH}}(\mathcal{P}_{\mathrm{FH}}^\star) + P^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}}^\star)} \\ &= \max_{\{\mathcal{P}_{\mathrm{FH}}, \mathcal{P}_{\mathrm{AC}}\}} \frac{\sum_{n=1}^{N}\sum_{k=1}^{K} R_{nk}^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}})}{P^{\mathrm{FH}}(\mathcal{P}_{\mathrm{FH}}) + P^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}})}, \end{aligned} \tag{10}$$

where $\mathcal{P}_{\mathrm{FH}}^\star$ and $\mathcal{P}_{\mathrm{AC}}^\star$ are the power allocation policy corresponding to $\eta^\star$. Based on the above, we have the following theorem.

Theorem 1: The maximum EE $\eta^\star$ is achieved if and only if

$$\begin{aligned} &\max_{\{\mathcal{P}_{\mathrm{FH}}, \mathcal{P}_{\mathrm{AC}}\}} \sum_{n=1}^{N}\sum_{k=1}^{K} R_{nk}^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}}) - \eta^\star(P^{\mathrm{FH}}(\mathcal{P}_{\mathrm{FH}}) + P^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}})) \\ &= \sum_{n=1}^{N}\sum_{k=1}^{K} R_{nk}^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}}^\star) - \eta^\star(P^{\mathrm{FH}}(\mathcal{P}_{\mathrm{FH}}^\star) + P^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}}^\star)) \\ &= 0 \end{aligned} \tag{11}$$

Proof: Refer to [21]. ∎

Based on Theorem 1, the original EE problem can be transformed into the following one with parameter η

$$\max_{\{\mathcal{P}_{\mathrm{FH}}, \mathcal{P}_{\mathrm{AC}}, \eta\}} \sum_{n=1}^{N}\sum_{k=1}^{K} R_{nk}^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}}) - \eta(P^{\mathrm{FH}}(\mathcal{P}_{\mathrm{FH}}) + P^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}})) \tag{12a}$$

$$\text{s.t. (9b) – (9f).} \tag{12b}$$

Algorithm 1: Dinkelbach-Based Iterative Algorithm

Initialize the maximum number of iterations L_{max}, the maximum tolerate ε, the maximum EE $\eta = 0$, and the iteration index $t = 0$.
repeat
  Solve the problem (12) for a given η and obtain the power allocation $\{\mathcal{P}_{\mathrm{FH}}, \mathcal{P}_{\mathrm{AC}}\}$.
  Compute $\varepsilon^\star = \sum_{n=1}^{N}\sum_{k=1}^{K} R_{nk}^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}}) - \eta(P^{\mathrm{FH}}(\mathcal{P}_{\mathrm{FH}}) + P^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}}))$.
  if $\varepsilon^\star < \varepsilon$ **then**
    Covergence=**true**.
    return $\{\mathcal{P}_{\mathrm{FH}}^\star, \mathcal{P}_{\mathrm{AC}}^\star\} = \{\mathcal{P}_{\mathrm{FH}}, \mathcal{P}_{\mathrm{AC}}\}$ and $\eta^\star = \frac{\sum_{n=1}^{N}\sum_{k=1}^{K} R_{nk}^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}})}{P^{\mathrm{FH}}(\mathcal{P}_{\mathrm{FH}}) + P^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}})}$.
  else
    Set $\eta = \frac{\sum_{n=1}^{N}\sum_{k=1}^{K} R_{nk}^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}})}{P^{\mathrm{FH}}(\mathcal{P}_{\mathrm{FH}}) + P^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}})}$ and $t = t + 1$.
    Convergence=**false**.
  end if
until Convergence=**true** or $t = L_{max}$;

To obtain the optimal η, Dinkelbach-based iterative method [21] is usually used by setting the initial value of η, and then iterating solve (12). We summarize the above iterative method in Algorithm 1. Next, we need to solve the optimization problem (12) for a given η, which is still difficult to solve due to the non-convex objection function (12a) and the non-convex fronthaul capacity constraint (9e).

B. The Transformation of Fronthaul Capacity Constraints

It is well known that ℓ_0-norm is the number of nonzero entries in a vector. Furthermore, the ℓ_0-norm optimization problem can be usually approximated by a reweighed ℓ_1-norm as follows [22]

$$\|\mathbf{A}\|_0 \approx \sum_i \alpha_i |a_i|, \tag{13}$$

where a_i is the ith component in the vector $\mathbf{A}$ and α_i denotes the weighted coefficient. Accordingly, the minimization problem of $\|\mathbf{A}\|_0$ can be transformed to minimize $\sum_i \alpha_i |a_i|$ by setting proper weighted coefficient α_i. Based on this, we first indicate the following function

$$\mathbb{1}\left\{\|P_{nk}^{\mathrm{AC}}\|_2^2\right\} = \begin{cases} 0, & \text{if } \|P_{nk}^{\mathrm{AC}}\|_2^2 = 0, \\ 1, & \text{otherwise.} \end{cases} \tag{14}$$

As a result, the fronthaul capacity constraint (9e) can be rewrite as

$$\sum_{k=1}^{K} \mathbb{1}\left\{\|P_{nk}^{\mathrm{AC}}\|_2^2\right\} R_{nk}^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}}) \leq R_n^{\mathrm{FH}}(\mathcal{P}_{\mathrm{FH}}), n \in \{1, \cdots, N\}. \tag{15}$$

In addition, it is obvious that

$$\mathbb{1}\left\{\|P_{nk}^{\mathrm{AC}}\|_2^2\right\} = \left\|\|P_{nk}^{\mathrm{AC}}\|_2^2\right\|_0, \tag{16}$$

and we reformulate (15) as

$$\sum_{k=1}^{K} \alpha_{nk} P_{nk}^{\mathrm{AC}} \hat{R}_{nk}^{\mathrm{AC}}(\hat{\mathcal{P}}_{\mathrm{AC}}) - R_n^{\mathrm{FH}}(\mathcal{P}_{\mathrm{FH}}) \leq 0, n \in \{1, \cdots, N\}. \tag{17}$$

Algorithm 2: ℓ_1-norm-Based Iterative Algorithm

Initialize the maximum number of iterations L_{max}, the iteration index $t = 0$ and feasible power $\{\mathcal{P}_{\mathrm{FH}}^{(0)}, \mathcal{P}_{\mathrm{AC}}^{(0)}\}$.
repeat
  Compute $\alpha_{nk}^{(t)}$ and $\hat{R}_{nk}^{\mathrm{AC}}(\hat{\mathcal{P}}_{\mathrm{AC}}^{(t)})$, $t = t + 1$.
  Solve the problem (19) and obtain the power allocation $\{\mathcal{P}_{\mathrm{FH}}^{(t)}, \mathcal{P}_{\mathrm{AC}}^{(t)}\}$.
until $\{\mathcal{P}_{\mathrm{FH}}^{(t)}, \mathcal{P}_{\mathrm{AC}}^{(t)}\}$ converge or $t = L_{max}$;

where α_{nk} is a constant and $\hat{R}_{nk}^{\mathrm{AC}}(\hat{\mathcal{P}}_{\mathrm{AC}})$ is obtained from the previous iteration. Meanwhile, α_{nk} can be updated according to the following

$$\alpha_{nk} = \frac{1}{\hat{P}_{nk}^{\mathrm{AC}} + \epsilon}, \forall n, k, \tag{18}$$

where $\epsilon > 0$ is a small constant, while $\hat{P}_{nk}^{\mathrm{AC}}$ denotes optimal power for previous iteration. It is clear that (17) is a convex constraint, and thus we need to solve the following optimization problem

$$\max_{\{\mathcal{P}_{\mathrm{FH}}, \mathcal{P}_{\mathrm{AC}}\}} \sum_{n=1}^{N}\sum_{k=1}^{K} R_{nk}^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}}) - \eta^{\star}\left(P^{\mathrm{FH}}(\mathcal{P}_{\mathrm{FH}}) + P^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}})\right) \tag{19a}$$

$$\text{s.t. (9b)} - \text{(9d)}, \text{(9f)}, \text{(17)}. \tag{19b}$$

To obtain the effective power allocation, we need to iteratively solve (19) by updating parameter α_{nk}, and we summarize it in Algorithm 2.

C. WMMSE-Based Approach For Solving (19)

Although all constraints are convex, (19) is still a non-convex optimization problem due to the non-convex objective function. Next, we will equivalently transform (19a) into a convex one by MMSE scheme. Specifically, if the MMSE scheme is used to detect s_{nk} from received signal y_{nk}^{AC}, we have the following detection problem:

$$\theta_{nk}^{\star} = \arg \min_{\{\theta_{nk}\}} e_{nk}, \forall n, k, \tag{20}$$

where η_{nk} represents the receiver filter at the kth device served by the nth RRH, e_{nk} is its mean square error which can be expressed as

$$e_{nk} = \mathrm{E}\left\{\left|s_{nk} - \theta_{nk} y_{nk}^{\mathrm{AC}}\right|^2\right\}, \forall n, k. \tag{21}$$

Substituting (4) into e_{nk}, we can obtain

$$e_{nk} = 1 + |\theta_{nk}|^2\left(|g_{nnk}|^2 P_{nk}^{\mathrm{AC}} + \pi_{nk}\right) - 2\mathrm{Re}\left(\theta_{nk}\sqrt{P_{nk}^{\mathrm{AC}}} g_{nnk}\right). \tag{22}$$

Combining (22) and (20), we can obtain the optimal receiver filter $\theta_{nk}^{\star}$ as

$$\theta_{nk}^{\star} = \frac{\sqrt{P_{nk}^{\mathrm{AC}}} g_{nnk}^{*}}{|g_{nnk}|^2 P_{nk}^{\mathrm{AC}} + \pi_{nk}}, \forall n, k. \tag{23}$$

Based on (23), the MMSE $e_{nk}^{\star}$ can be expressed as

$$e_{nk}^{\star} = 1 - |g_{nnk}|^2 P_{nk}^{\mathrm{AC}}\left(|g_{nnk}|^2 P_{nk}^{\mathrm{AC}} + \pi_{nk}\right)^{-1}, \forall n, k. \tag{24}$$

The detailed derivation of (23) and (24) can refer to Appendix A. In addition, the following equation can be easily obtained through (5)

$$\begin{aligned}\left(1 + \frac{|g_{nnk}|^2 P_{nk}^{\mathrm{AC}}}{\pi_{nk}}\right)^{-1} &= \frac{\pi_{nk}}{|g_{nnk}|^2 P_{nk}^{\mathrm{AC}} + \pi_{nk}} \\ &= 1 - |g_{nnk}|^2 P_{nk}^{\mathrm{AC}}\left(|g_{nnk}|^2 P_{nk}^{\mathrm{AC}} + \pi_{nk}\right)^{-1}.\end{aligned} \tag{25}$$

Comparing (25) and (24), we find that they have the same expression for the right side of the equation, and thus we have the following

$$\left(1 + \frac{|g_{nnk}|^2 P_{nk}^{\mathrm{AC}}}{\pi_{nk}}\right)^{-1} = \arg \min_{\{\theta_{nk}\}} e_{nk}, \forall n, k. \tag{26}$$

Based on (26), the relationship between the throughput and MMSE can be expressed as

$$\begin{aligned}R_{nk}^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}}) &= W_{\mathrm{AC}}\log_2\left(1 + \frac{|g_{nnk}|^2 P_{nk}^{\mathrm{AC}}}{\pi_{nk}}\right) \\ &= -W_{\mathrm{AC}}\log_2\left(\min_{\{\theta_{nk}\}} e_{nk}\right) \\ &= \max_{\{\theta_{nk}\}}\left(-W_{\mathrm{AC}}\log_2(e_{nk})\right).\end{aligned} \tag{27}$$

After removing $\log_2$, (27) can be rewritten as

$$R_{nk}^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}}) = \max_{\{\theta_{nk}, d_{nk}\}} W_{\mathrm{AC}}\left(-\frac{d_{nk}e_{nk}}{\ln 2} + \log_2 d_{nk} + \frac{1}{\ln 2}\right), \forall n, k. \tag{28}$$

The detailed proof can refer to Appendix B.

Replacing $R_{nk}^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}})$ of objective function (19) with (28), we rewrite the optimization problem as follows

$$\max_{\{\mathcal{P}_{\mathrm{FH}}, \mathcal{P}_{\mathrm{AC}}, \theta_{nk}, d_{nk}\}} \sum_{n=1}^{N}\sum_{k=1}^{K} W_{\mathrm{AC}}\left(-\frac{d_{nk}e_{nk}}{\ln 2} + \log_2 d_{nk} + \frac{1}{\ln 2}\right) \tag{29a}$$

$$- \eta^{\star}\left(P^{\mathrm{FH}}(\mathcal{P}_{\mathrm{FH}}) + P^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}})\right) \tag{29b}$$

$$\text{s.t. } \pi_{nk}\omega_{nk} - |g_{nnk}|^2 P_{nk}^{\mathrm{AC}} \le \pi_{nk}, \forall n, k, \tag{29c}$$

$$\text{(9c)}, \text{(9d)}, \text{(9f)}, \text{(17)}. \tag{29d}$$

where $\omega_{nk} = 2^{\frac{R_{nk,\min}^{\mathrm{AC}}}{W_{\mathrm{AC}}}} - 1$. For the above optimization problem, WMMSE-based iterative algorithm can be used to solve it. First, for the given initial feasible power $\{\hat{\mathcal{P}}_{\mathrm{FH}}, \hat{\mathcal{P}}_{\mathrm{AC}}\}$, we can obtain the optimal $\theta_{nk}^{\star}$ given by (23) as

$$\theta_{nk}^{\star} = \frac{\sqrt{\hat{P}_{nk}^{\mathrm{AC}}} g_{nnk}^{*}}{|g_{nnk}|^2 \hat{P}_{nk}^{\mathrm{AC}} + \hat{\pi}_{nk}}, \forall n, k, \tag{30}$$

where $\hat{\pi}_{nk} = |g_{nnk}|^2 \sum_{i=1}^{k-1} \hat{P}_{ni}^{\mathrm{AC}} + \sum_{j\neq n}^{N} |g_{jnk}|^2 \sum_{i=1}^{K} \hat{P}_{ji}^{\mathrm{AC}} + W_{\mathrm{AC}} N_0$. Meanwhile, the optimal MMSE $e_{nk}^{\star}$ can be calculated by (24)

$$e_{nk}^{\star} = 1 - |g_{nnk}|^2 \hat{P}_{nk}^{\mathrm{AC}}\left(|g_{nnk}|^2 \hat{P}_{nk}^{\mathrm{AC}} + \hat{\pi}_{nk}\right)^{-1}, \forall n, k. \tag{31}$$

Based on (31), the optimal $d_{nk}^{\star}$ can be expressed as

$$d_{nk}^{\star} = \frac{1}{e_{nk}^{\star}}, \forall n, k. \tag{32}$$

Then, we only need to solve the following optimization problem for the next iteration

$$\min_{\{\mathcal{P}_{\mathrm{FH}},\mathcal{P}_{\mathrm{AC}}\}} \sum_{n=1}^{N}\sum_{k=1}^{K}\frac{W_{\mathrm{AC}}d_{nk}^{\star}e_{nk}}{\ln 2}-\eta^{\star}\xi\left(\sum_{n=1}^{N}\sum_{k=1}^{K}P_{nk}^{\mathrm{AC}}+\sum_{n=1}^{N}P_{n}^{\mathrm{FH}}\right) \tag{33a}$$

$$\text{s.t. } \pi_{nk}\omega_{nk}-|g_{nnk}|^2P_{nk}^{\mathrm{AC}}\le\pi_{nk},\forall n,k, \tag{33b}$$

$$\sum_{k=1}^{K}\alpha_{nk}P_{nk}^{\mathrm{AC}}\hat{R}_{nk}^{\mathrm{AC}}(\hat{\mathcal{P}}_{\mathrm{AC}})-R_{n}^{\mathrm{FH}}(\mathcal{P}_{\mathrm{FH}})\le 0, \tag{33c}$$

$$\sum_{n=1}^{N}P_{n}^{\mathrm{FH}}\le P_{\max}^{\mathrm{FH}},\sum_{k=1}^{K}P_{nk}^{\mathrm{AC}}\le P_{n,\max}^{\mathrm{AC}}, \tag{33d}$$

$$P_{n}^{\mathrm{FH}}\ge 0,P_{nk}^{\mathrm{AC}}\ge 0,\forall n,k. \tag{33e}$$

It is clear that (33) is a convex optimization problem. Therefore, we can obtain the optimal solution of (33) by solving its dual problem due to the zero-gap solution between the original problem and its dual problem [23]. First, we define the Lagrange function as follows

$$\begin{aligned}
&L(\mathcal{P}_{\mathrm{FH}},\mathcal{P}_{\mathrm{AC}},\mathcal{U},\mathcal{V},\mathcal{T},z)\\
&=\sum_{n=1}^{N}\sum_{k=1}^{K}\frac{W_{\mathrm{AC}}d_{nk}^{\star}e_{nk}}{\ln 2}-\eta^{\star}\xi\left(\sum_{n=1}^{N}\sum_{k=1}^{K}P_{nk}^{\mathrm{AC}}+\sum_{n=1}^{N}P_{n}^{\mathrm{FH}}\right)\\
&+\sum_{n=1}^{N}\sum_{k}^{K}u_{nk}\left(\pi_{nk}\omega_{nk}-|g_{nnk}|^2P_{nk}^{\mathrm{AC}}-\pi_{nk}\right)\\
&+\sum_{n=1}^{N}v_{n}\left(\sum_{k=1}^{K}\alpha_{nk}P_{nk}^{\mathrm{AC}}\hat{R}_{nk}^{\mathrm{AC}}(\hat{\mathcal{P}}_{\mathrm{AC}})-R_{n}^{\mathrm{FH}}(\mathcal{P}_{\mathrm{FH}})\right)\\
&+\sum_{n=1}^{N}t_{n}\left(\sum_{k=1}^{K}P_{nk}^{\mathrm{AC}}-P_{n,\max}^{\mathrm{AC}}\right)+z\left(\sum_{n=1}^{N}P_{n}^{\mathrm{FH}}-P_{\max}^{\mathrm{FH}}\right),
\end{aligned} \tag{34}$$

where $\mathcal{U}=\{u_{nk}\}_{N\times K}$, $\mathcal{V}=\{v_n\}_{N\times 1}$,$\mathcal{T}=\{t_n\}_{N\times 1}$ and z are the Lagrange variables associated with (33b)-(33d), respectively. Then, the Lagrange dual function can be expressed as

$$g(\mathcal{U},\mathcal{V},\mathcal{T},z)=\min_{\{\mathcal{P}_{\mathrm{FH}},\mathcal{P}_{\mathrm{AC}}\}}L(\mathcal{P}_{\mathrm{FH}},\mathcal{P}_{\mathrm{AC}},\mathcal{U},\mathcal{V},\mathcal{T},z), \tag{35}$$

and the dual optimization problem becomes

$$\max_{\{\mathcal{U},\mathcal{V},\mathcal{T},z\}} g(\mathcal{U},\mathcal{V},\mathcal{T},z), \tag{36a}$$

$$\text{s.t. } \mathcal{U}\ge 0,\mathcal{V}\ge 0,\mathcal{T}\ge 0,z\ge 0. \tag{36b}$$

Since the above dual function is convex, we maximize $g(\mathcal{U},\mathcal{V},\mathcal{T},z)$ by subgradient method [23], and the dual variables can be updated as follows:

$$\begin{aligned}
u_{nk}(o+1)&=\left[u_{nk}(o)+\psi_1(o)\left(\pi_{nk}\omega_{nk}-|g_{nnk}|^2P_{nk}^{\mathrm{AC}}-\pi_{nk}\right)\right]^{+},\\
v_{n}(o+1)&=\left[v_{n}(o)+\psi_2(o)\left(\sum_{k=1}^{K}\alpha_{nk}P_{nk}^{\mathrm{AC}}\hat{R}_{nk}^{\mathrm{AC}}(\hat{\mathcal{P}}_{\mathrm{AC}})-R_{n}^{\mathrm{FH}}(\mathcal{P}_{\mathrm{FH}})\right)\right]^{+},\\
t_{n}(o+1)&=\left[t_{n}(o)+\psi_3(o)\left(\sum_{k=1}^{K}P_{nk}^{\mathrm{AC}}-P_{n,\max}^{\mathrm{AC}}\right)\right]^{+},\\
z(o+1)&=\left[z(o)+\psi_4(o)\left(\sum_{n=1}^{N}P_{n}^{\mathrm{FH}}-P_{\max}^{\mathrm{FH}}\right)\right]^{+},
\end{aligned} \tag{37}$$

Algorithm 3: WMMSE-Based Iterative Algorithm

Initialize feasible $\{\mathcal{P}_{\mathrm{FH}}^{(0)},\mathcal{P}_{\mathrm{AC}}^{(0)}\}$, t=0.
repeat
  Compute $\{\theta_{nk}^{\star(t)}\}$ according to (30).
  Compute $\{d_{nk}^{\star(t)}\}$ according to (32).
  $t=t+1$.
  Compute $\{\hat{\mathcal{P}}_{\mathrm{FH}}^{(t)},\hat{\mathcal{P}}_{\mathrm{AC}}^{(t)}\}$ according to (38).
until $\{\hat{\mathcal{P}}_{\mathrm{FH}}^{(t)},\hat{\mathcal{P}}_{\mathrm{AC}}^{(t)}\}$ converges;

where $\psi_n(o)$ is the positive step size at the oth iteration. Then, for fixed dual variables, we can obtain the optimal power allocation by Karush-Kuhn-Tucker (KKT) conditions as

$$\begin{aligned}
P_{nk}^{\mathrm{AC}}&=\left(\frac{W_{\mathrm{FH}}d_{nk}^{\star}\mathrm{Re}\left(\theta_{nk}g_{nnk}\right)}{\Xi_{nk}\ln 2}\right)^2,\\
P_{n}^{\mathrm{FH}}&=\left[\frac{v_nW_{\mathrm{FH}}}{\ln 2}-\frac{W_{\mathrm{FH}}N_0}{\|\mathbf{h}_n\mathbf{v}_n\|^2}\right]^{+},
\end{aligned} \tag{38}$$

where $\Xi_{nk}=\frac{W_{\mathrm{AC}}d_{nk}^{\star}}{\ln 2}\left[|\theta_{nk}|^2|g_{nnk}|^2+\sum_{i=k+1}^{K}|\theta_{ni}|^2|g_{nni}|^2\right]+\frac{W_{\mathrm{AC}}d_{nk}^{\star}}{\ln 2}\left[\sum_{j\neq n}^{N}\sum_{i=1}^{K}|\theta_{ji}|^2|g_{nji}|^2\right]-\eta^{\star}\xi-u_{nk}|g_{nnk}|^2+\sum_{i=k+1}^{K}u_{ni}(\omega_{ni}-1)|g_{nni}|^2+\sum_{j\neq n}^{N}\sum_{i=1}^{K}u_{ji}(\omega_{ji}-1)|g_{nji}|^2+v_n\alpha_{nk}\hat{R}_{nk}^{\mathrm{AC}}(\hat{P}_{AC})+t_n$. The detailed proof can refer to Appendix C. Based on the above analysis, the optimization scheme to solve problem (19) is summarized as the Algorithm 3.

From the Algorithm 3, it is clear that the optimal power allocation can be obtained due to the convex optimization problem (33) in each iteration. Meanwhile, the optimal $\theta_{nk}^{\star}$ and $d_{nk}^{\star}$ can be updated at each iteration. To this end, iteratively updating the above parameters will increases or at least maintain the value of the objective function (29) [24]. Therefore, Algorithm 3 will converge to a stationary solution due to the limited transmit power.

D. Algorithm Complexity Analysis

First, we summarize the overall algorithm to solve the optimization problem (9). We need to set the initial feasible power $\{\mathcal{P}_{\mathrm{FH}},\mathcal{P}_{\mathrm{AC}}\}$ and η, and thus compute α_{nk} and $R_{nk}^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}})$. After that, Algorithm 3 is carried out. Then, we execute Algorithm 2 and then Algorithm 1 until $\{\mathcal{P}_{\mathrm{FH}},\mathcal{P}_{\mathrm{AC}}\}$ and η convergence.

Next, we analyze the complexity of the overall algorithm. Since the number of dual variables is $NK+2N+1$, the optimal dual variables is obtained by using subgradient updated method with the complexity $O(|NK+2N+1|^2)$. As a result, the complexity of Algorithm 3 is $O(e_3|NK+2N+1|^2NK)$, where e_3 is the maximum iteration times of Algorithm 3. We assume the maximum iteration times of Algorithms 1 and 2 are e_1 and e_2, respectively. The complexity of the overall algorithm is $O(e_1e_2e_3|NK+2N+1|^2NK)$.

IV. Simulation Results

In this section, simulation results are provided to show the throughput and EE of the system under the proposed algorithm. For comparison, we also provide the results for OMA scheme, where the time division multiple access (TDMA) is

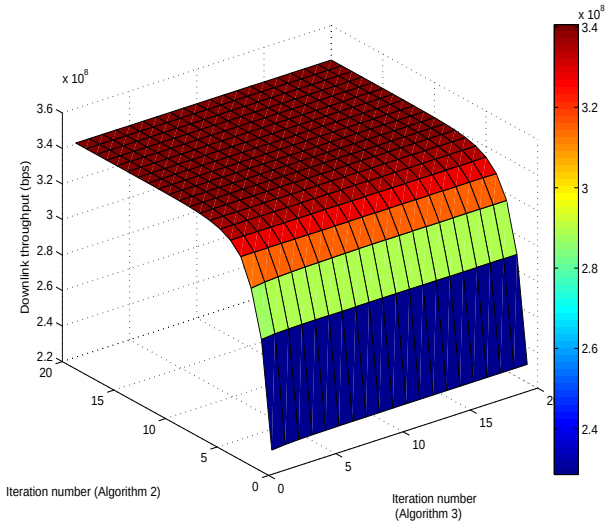

Fig. 2: The convergence of Algorithms 2 and 3.

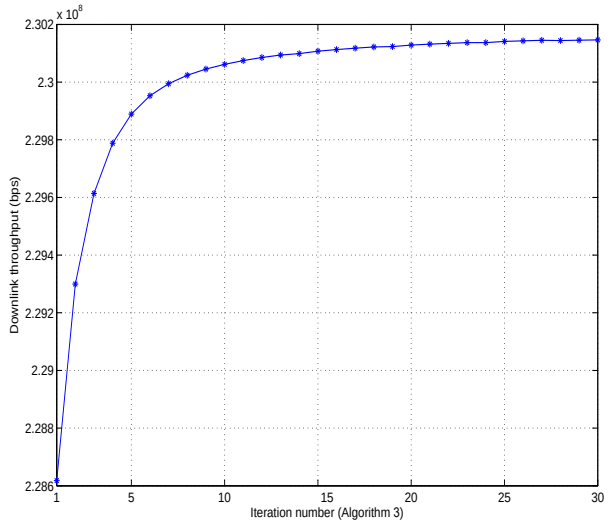

Fig. 3: The convergence of Algorithm 3.

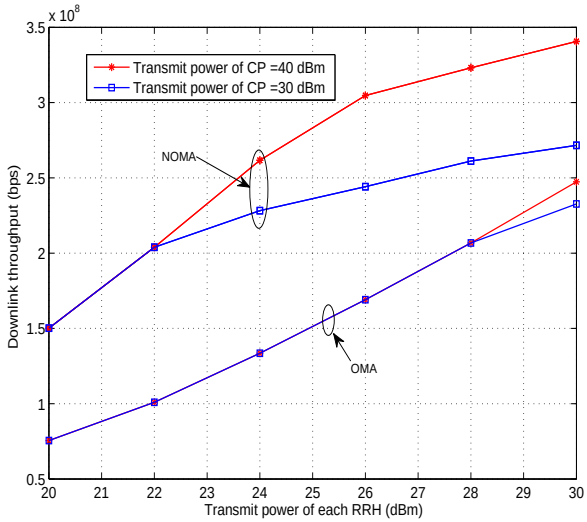

Fig. 4: The downlink throughput versus transmit power of each RRH.

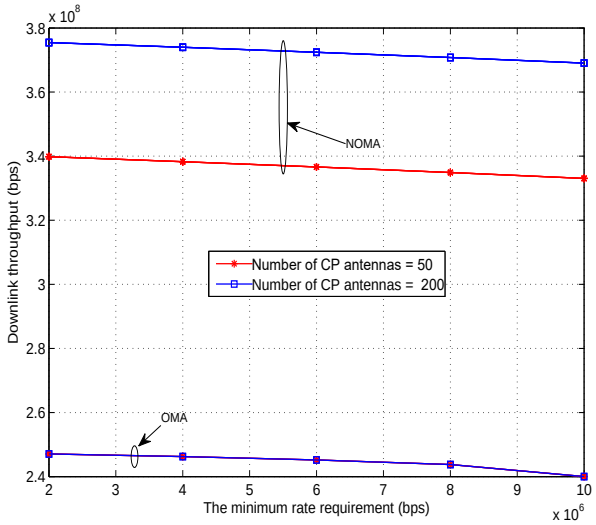

Fig. 5: The downlink throughput versus the minimum rate requirement.

adopted. Specifically, the RRH transmits the downlink data to device with TDMA scheme. We consider N =3 RRHs as shown in Fig. 1, and each RRH serves K = 2 devices. The coverage radius of each RRH is 200 meters, and the distance between the RRH and the CP is 1500 meters. The mmWave channel is centered at 73 GHz with a bandwidth of W_{AC}=50 MHz and the path loss is modeled as $69.7 + 24\log_{10}(d)$ dB, where d denotes the distance (meter) [8], [25]. The wireless fronthaul channel is centered at a frequency of 2 GHz with a bandwidth of W_{FH} = 10 MHz, and the path loss is modeled as $120 + 30\log_{10}(d)$ dB, where d denotes the distance (kilometer) [26]. The circuit power consumption P_c^{FH} = 10 W and $P_{c,n}^{\mathrm{AC}}$=1 W. The noise power spectral density is -174 dBm/Hz, and the reciprocal of drain efficiency of power amplifier ξ = 0.38. ϵ is set as 10^{-8}. Other related parameters will be presented in discussion.

A. *Throughput of the System*

In this subsection, we evaluate the throughput of the system, where η is set as 0. Fig. 2 shows the convergence of Algorithms 2 and 3. Here, we set $P_{\max}^{\mathrm{FH}}$ = 40 dBm, $P_{n,\max}^{\mathrm{AC}}$ = 30 dBm, M = 50, and $R_{nk,\min}^{\mathrm{AC}}$ = 10^6 bps ($\forall n, k$). It is clear that Algorithm 2 (i.e., outer iteration) tends to converge after 10 iterations. In addition, to clearly show the convergence of Algorithm 3, we present Fig. 3 when the first iteration is executed in Algorithm 2. It can be found that the Algorithm 3 tends to converge after 25 iteration.

Fig. 4 shows the downlink throughput versus transmit power of each RRH under different transmit power of CP. We set the number of CP antennas M = 50. It can be observed that the downlink throughput increases with $P_{n,\max}^{\mathrm{AC}}$. It is easy to understand that higher transmit power contributes to more

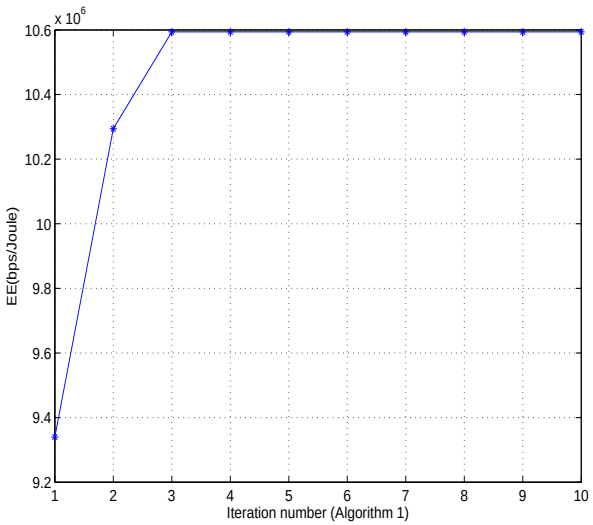

Fig. 6: The convergence of Algorithm 1.

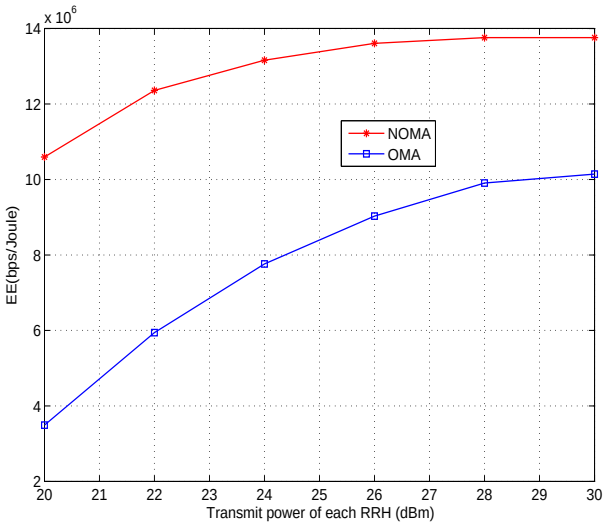

Fig. 7: EE versus transmit power of each RRH.

throughput. In addition, when the transmit power of CP is 40 dBm, the throughput is higher for an higher transmit power of each RRH. However, when the transmit power of each RRH is small (e.g., 20~22 dBm), the throughput is the same. This is because although the higher transmit power of CP can provide larger fronthaul capacity, the throughput of the access link is limited. On the other hand, we can find that the throughput under NOMA scheme is higher than that under OMA.

Fig. 5 presents the downlink throughput versus the minimum rate requirement of devices under different number of CP antennas. Here, we set $P_{\max}^{\mathrm{FH}} = 40$ dBm, $P_{n,\max}^{\mathrm{AC}} = 30$ dBm. It is clearly observed that the throughput decreases with the minimum rate requirement. This is because that RRHs must transmit proper power to guarantee the minimum rate requirement for each device so as to sacrifice some total throughput. In addition, we can find that more number of CP antennas can obtain higher throughput under NOMA. It is easy to understand that the higher antenna gain can be obtained for more antennas so as to improve the fronthaul capacity, which can provide higher throughput for the access link. For the NOMA scheme, we find that the throughput is the same for different number of CP antennas. It is because that the throughput of the access link is limited. In addition, it is clear that the throughput under the NOMA scheme is higher than that under the OMA scheme.

B. EE of the System

In this subsection, we evaluate the EE of the system. Fig. 6 shows the convergence of Algorithm 1. Here, we set $P_{\max}^{\mathrm{FH}} = 30$ dBm, $P_{n,\max}^{\mathrm{AC}} = 20$ dBm, $M = 50$, and $R_{nk,\min}^{\mathrm{AC}} = 10^6$ bps $(\forall n, k)$. We can find that the EE rapidly increases and tends to stabilize after 3 iterations. Here, we do not show the convergence of Algorithms 2 and 3, and they have the similar characteristic with Figs. 2 and 3.

Fig.7 shows the EE versus the transmit power of each RRH. Here, we set $P_{\max}^{\mathrm{FH}} = 30$ dBm, $M = 50$, and $R_{nk,\min}^{\mathrm{AC}} = 10^6$ bps $(\forall n, k)$. It can be observed that the EE first increases and then

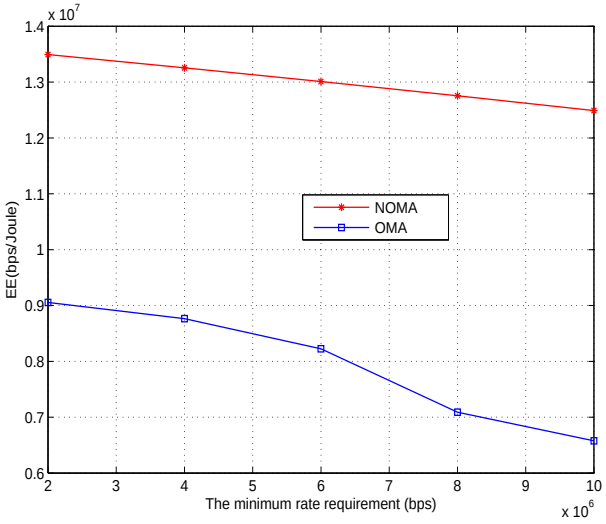

Fig. 8: EE versus the minimum rate requirement.

tends to stabilize with the transmit power of each RRH. It is because that although the higher transmit power of the RRHs can provide higher throughout, it need to consume more power to improve the unit throughput. Based on this, the increase of the EE will be slower and slower with the transmit power as shown in Fig. 7. In addition, the EE under the OMA has the similar trend with the transmit power of the RRHs. One the other hand, the EE under the NOMA scheme is higher than that under the OMA scheme.

Fig. 8 presents the EE versus the minimum rate requirement. Here, we set $P_{\max}^{\mathrm{FH}} = 30$ dBm, $P_{n,\max}^{\mathrm{AC}} = 26$ dBm, and $M = 50$. We can clearly find that the EE decreases with the minimum rate requirement. In fact, we have found the similar results at Fig. 5, namely the throughput decreases with the minimum rate requirement. Therefore, they have the same reason for it, i.e., the RRHs must guarantee the devices' rate requirement at the sacrifice of the EE. In addition, it can be easily found that

the EE is higher under the NOMA scheme that that under the OMA.

V. Conclusions

In this paper, we investigated the throughput and EE in a NOMA-based CRAN, where each RRH serves devices with NOMA scheme. Based on this, we formed the power optimization problem of the CP and RRHs to maximize the EE of the system while satisfying the fronthaul capacity and transmit power constraints of the CP and RRHs. Due to the difficulty of the formulated problem to directly solve, we first transformed the original fractional objective function into a subtractive one by bringing the parameter. The optimal parameter was found by Dinkelbach-based iterative algorithm. Next, we proposed a two-loop iterative algorithm to solve the parameter-based optimization problem. For the outer iteration, we applied the ℓ_1-norm-based technique transform the non-convex fronthaul capacity constraint into the convex one. For the inner iteration, the WMMSE-based approach was used to deal with the finial problem. Finally, we found that the throughput and EE of the system under the proposed NOMA scheme outperform that under the conventional OMA scheme.

Appendix A

According to (22), we have

$$\begin{aligned} e_{nk} &= 1+|\theta_{nk}|^2\left(|g_{nnk}|^2P_{nk}^{\mathrm{AC}}+\pi_{nk}\right)-2\mathrm{Re}\left(\theta_{nk}\sqrt{P_{nk}^{\mathrm{AC}}}g_{nnk}\right)\\ &=\left|1-\theta_{nk}\sqrt{P_{nk}^{\mathrm{AC}}}g_{nnk}\right|^2+|\theta_{nk}|^2|g_{nnk}|^2\sum_{i=1}^{k-1}P_{ni}^{\mathrm{AC}}\\ &\quad+|\theta_{nk}|^2\sum_{j\neq n}^{N}|g_{jnk}|^2\sum_{i=1}^{K}P_{ji}^{\mathrm{AC}}+|\theta_{nk}|^2W_{\mathrm{AC}}N_0, \end{aligned} \tag{39}$$

To minimize e_{nk} by optimizing wighted coefficient θ_{nk}, i.e., solving (20), we can take the derivative of e_{nk} for θ_{nk}, namely

$$\frac{\partial e_{nk}}{\partial \theta_{nk}}=\theta_{nk}^{*}\left(|g_{nnk}|^2P_{nk}^{\mathrm{AC}}+\pi_{nk}\right)-\sqrt{P_{nk}^{\mathrm{AC}}}g_{nnk}=0. \tag{40}$$

Therefore, the optimal weighted coefficient $\theta_{nk}^{\star}$ can be calculated as

$$\theta_{nk}^{\star}=\frac{\sqrt{P_{nk}^{\mathrm{AC}}}g_{nnk}^{*}}{|g_{nnk}|^2P_{nk}^{\mathrm{AC}}+\pi_{nk}},\forall n,k. \tag{41}$$

Combing (41) and (22), we have

$$\begin{aligned} e_{nk}^{\star} &= 1+|\theta_{nk}^{\star}|^2\left(|g_{nnk}|^2P_{nk}^{\mathrm{AC}}+\pi_{nk}\right)-2\mathrm{Re}\left(\theta_{nk}^{\star}\sqrt{P_{nk}^{\mathrm{AC}}}g_{nnk}\right)\\ &=1+\left|\sqrt{P_{nk}^{\mathrm{AC}}}g_{nnk}^{*}\left(|g_{nnk}|^2P_{nk}^{\mathrm{AC}}+\pi_{nk}\right)^{-1}\right|^2\left(|g_{nnk}|^2P_{nk}^{\mathrm{AC}}+\pi_{nk}\right)\\ &\quad-2\mathrm{Re}\left(\sqrt{P_{nk}^{\mathrm{AC}}}g_{nnk}^{*}\left(|g_{nnk}|^2P_{nk}^{\mathrm{AC}}+\pi_{nk}\right)^{-1}\sqrt{P_{nk}^{\mathrm{AC}}}g_{nnk}\right)\\ &=1+|g_{nnk}|^2P_{nk}^{\mathrm{AC}}\left(|g_{nnk}|^2P_{nk}^{\mathrm{AC}}+\pi_{nk}\right)^{-1}\\ &\quad-2|g_{nnk}|^2P_{nk}^{\mathrm{AC}}\left(|g_{nnk}|^2P_{nk}^{\mathrm{AC}}+\pi_{nk}\right)^{-1}\\ &=1-|g_{nnk}|^2P_{nk}^{\mathrm{AC}}\left(|g_{nnk}|^2P_{nk}^{\mathrm{AC}}+\pi_{nk}\right)^{-1}. \end{aligned} \tag{42}$$

We finish the derivation.

Appendix B

First, we define the following function

$$f(d)=-\frac{bd}{\ln 2}+\log_2 d+\frac{1}{\ln 2}, \tag{43}$$

where b is a positive real number. Based on this, we formulate the following optimization problem.

$$\max_{\{d>0\}}\ f(d). \tag{44}$$

Since function $f(d)$ is concave with respect to d, we can directly solving $\frac{\partial f(d)}{\partial d}\Big|_{d=d^{\star}}=0$ to obtain the optimal $d^{\star}$ for maximizing $f(d)$, namely

$$\frac{\partial f(d)}{\partial d}\Big|_{d=d^{\star}}=-\frac{b}{\ln 2}+\frac{1}{d\ln 2}=0\quad\Rightarrow d^{\star}=\frac{1}{b}, \tag{45}$$

Substituting $d^{\star}$ into $f(d)$, we have

$$\max_{\{d>0\}}\ f(d)=-\log_2 b. \tag{46}$$

By replacing b with e_{nk} and combining (27), we have

$$\begin{aligned} R_{nk}^{\mathrm{AC}}(\mathcal{P}_{\mathrm{AC}}) &=\max_{\{\theta_{nk}\}}\left(-W_{\mathrm{AC}}\log_2(e_{nk})\right)\\ &=\max_{\{\theta_{nk}\}}\left(W_{\mathrm{AC}}\max_{\{d_{nk}\}}\left(-\frac{d_{nk}e_{nk}}{\ln 2}+\log_2 d_{nk}+\frac{1}{\ln 2}\right)\right)\\ &=\max_{\{\theta_{nk},d_{nk}\}}W_{\mathrm{AC}}\left(-\frac{d_{nk}e_{nk}}{\mathrm{nk}2}+\log_2 d_{nk}+\frac{1}{\ln 2}\right),\forall n,k. \end{aligned} \tag{47}$$

We finish the proof.

Appendix C

The KKT conditions of (33) can be formulated as

$$\begin{aligned} \frac{\partial L}{\partial P_{nk}^{\mathrm{AC}}} &= \frac{W_{\mathrm{AC}}d_{nk}^{\star}}{\ln 2}\left[|\theta_{nk}|^2|g_{nnk}|^2-\mathrm{Re}\left(\theta_{nk}g_{nnk}\left(P_{nk}^{\mathrm{AC}}\right)^{-\frac{1}{2}}\right)\right]\\ &\quad+\frac{W_{\mathrm{AC}}d_{nk}^{\star}}{\ln 2}\left[\sum_{i=k+1}^{K}|\theta_{ni}|^2|g_{nni}|^2+\sum_{j\neq n}^{N}\sum_{i=1}^{K}|\theta_{ji}|^2|g_{nji}|^2\right]\\ &\quad-\eta^{\star}\xi-u_{nk}|g_{nnk}|^2+\sum_{i=k+1}^{K}u_{ni}(\omega_{ni}-1)|g_{nni}|^2\\ &\quad+\sum_{j\neq n}^{N}\sum_{i=1}^{K}u_{ji}(\omega_{ji}-1)|g_{nji}|^2+v_n\alpha_{nk}\hat{R}_{nk}^{\mathrm{AC}}(\hat{P}_{AC})+t_n\\ &\quad=0,\\ \frac{\partial L}{\partial P_n^{\mathrm{FH}}} &= -\eta^{\star}\xi-\frac{1}{2}\frac{v_nW_{\mathrm{FH}}\|\mathbf{h}_n\mathbf{v}_n\|^2}{W_{\mathrm{FH}}N_0+\|\mathbf{h}_n\mathbf{v}_n\|^2P_n^{\mathrm{FH}}}+z=0,\\ u_{nk}&\left(\pi_{nk}\omega_{nk}-|g_{nnk}|^2P_{nk}^{\mathrm{AC}}-\pi_{nk}\right)=0,\\ v_n&\left(\sum_{k=1}^{K}\alpha_{nk}P_{nk}^{\mathrm{AC}}\hat{R}_{nk}^{\mathrm{AC}}(\hat{\mathcal{P}}_{\mathrm{AC}})-R_n^{\mathrm{FH}}(\mathcal{P}_{\mathrm{FH}})\right)=0,\\ t_n&\left(\sum_{k=1}^{K}P_{nk}^{\mathrm{AC}}-P_{n,\max}^{\mathrm{AC}}\right)=0,\\ z&\left(\sum_{n=1}^{N}P_n^{\mathrm{FH}}-P_{\max}^{\mathrm{FH}}\right)=0. \end{aligned} \tag{48}$$

From (48), we can obtain the optimal power allocation as

$$P_{nk}^{\mathrm{AC}} = \left(\frac{W_{\mathrm{FH}} d_{nk}^{\star} \mathrm{Re}\left(\theta_{nk} g_{nnk}\right)}{\Xi_{nk} \ln 2} \right)^2, \\ P_{n}^{\mathrm{FH}} = \left[\frac{v_n W_{\mathrm{FH}}}{\ln 2} - \frac{W_{\mathrm{FH}} N_0}{\|\mathbf{h}_n \mathbf{v}_n\|^2} \right]^{+}, \tag{49}$$

where $\Xi_{nk} = \frac{W_{\mathrm{AC}} d_{nk}^{\star}}{\ln 2} \left[|\theta_{nk}|^2 |g_{nnk}|^2 + \sum_{i=k+1}^{K} |\theta_{ni}|^2 |g_{nni}|^2 \right] + \frac{W_{\mathrm{AC}} d_{nk}^{\star}}{\ln 2} \left[\sum_{j \neq n}^{N} \sum_{i=1}^{K} |\theta_{ji}|^2 |g_{njl}|^2 \right] - \eta^{\star} \xi - u_{nk} |g_{nnk}|^2 + \sum_{i=k+1}^{K} u_{ni} (\omega_{ni} - 1) |g_{nni}|^2 + \sum_{j \neq n}^{N} \sum_{i=1}^{K} u_{ji} (\omega_{ji} - 1) |g_{njl}|^2 + v_n \alpha_{nk} \hat{R}_{nk}^{\mathrm{AC}} (\hat{P}_{AC}) + t_n.$

References

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

[2] J. Huang, Y. Meng, X. Gong, Y. Liu, and Q. Duan, "A novel deployment scheme for green internet of things," *IEEE Internet Things J.*, vol. 1, no. 2, pp. 196205, Apr. 2014.

[3] J. Huang, Q. Duan, C. C. Xing, and H. Wang, "Topology control for building a large-scale and energy-efficient internet of things," *IEEE Wireless Commun.*, vol. 24, no. 1, pp. 67-73, Feb. 2017.

[4] M. Shirvanimoghaddam, M. Dohler and S. J. Johnson, "Massive non-orthogonal multiple access for cellular IoT: Potentials and limitations," *IEEE Commun. Magazine*, vol. 55, no. 9, pp. 55-61, 2017.

[5] Z. Ding et al., "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Commun. Magazine*, vol. 55, no. 2, pp. 185-191, Feb. 2017.

[6] C. Chen, W. Cai, X. Cheng, L. Yang and Y. Jin, "Low complexity beamforming and device selection schemes for 5G MIMO-NOMA systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2708-2722, Dec. 2017.

[7] J. G. Andrews et al., "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065-1082, Jun. 2014.

[8] R. G. Stephen and R. Zhang, "Joint millimeter-wave fronthaul and OFDMA resource allocation in ultra-dense CRAN," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1411-1423, Mar. 2017.

[9] V. N. Ha, L. B. Le and N. D. Dao, "Coordinated multipoint transmission design for cloud-RANs with limited fronthaul capacity constraints," *IEEE Trans. Veh. Technol.*, vol. 65, no. 9, pp. 7432-7447, Sep. 2016.

[10] S.-H. Park, O. Simeone, O. Sahin, and S. Shamai, "Joint decompression and decoding for cloud radio access networks," *IEEE Signal Processing Lett.*, vol. 20, no. 5, pp. 503-506, May 2013

[19] X. Gu, X. Ji, Z. Ding, W. Wu and M. Peng, "Outage probability analysis of non-orthogonal multiple access in cloud radio access networks," *IEEE Commun. Lett.*, vol. 22, no. 1, pp. 149-152, Jan. 2018.

[11] X. Huang, G. Xue, R. Yu and S. Leng, "Joint scheduling and beamforming coordination in cloud radio access networks with QoS guarantees," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5449-5460, Jul. 2016.

[12] W. Wang, V. K. N. Lau and M. Peng, "Delay-aware uplink fronthaul allocation in cloud radio access networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4275-4287, Jul. 2017.

[13] B. Dai and W. Yu, "Sparse beamforming and device-centric clustering for downlink cloud radio access network," *IEEE Access*, vol. 2, pp. 1326-1339, 2014.

[14] J. Liu, S. Zhou, J. Gong, Z. Niu and S. Xu, "Statistical multiplexing gain analysis of heterogeneous virtual base station pools in cloud radio access networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5681-5694, Aug. 2016.

[15] S. Gu, Z. Li, C. Wu and H. Zhang, "Virtualized resource sharing in cloud radio access networks through truthful mechanisms," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1105-1118, Mar. 2017.

[16] M. Peng, Y. Wang, T. Dang and Z. Yan, "Cost-efficient resource allocation in cloud radio access networks with heterogeneous fronthaul expenditures," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4626-4638, Jul. 2017.

[17] Q. Liu, G. Wu, Y. Guo, Y. Zhang and S. Hu, "Energy efficient resource allocation for control data separated heterogeneous-CRAN," IEEE Global Communications Conference, Washington, DC, 2016, pp. 1-6.

[18] B. Hu, C. Hua, C. Chen and X. Guan, "Joint beamformer design for wireless fronthaul and access links in C-RANs," *IEEE Trans. Wireless Commun.*, vol. PP, no. 99, pp. 1-1.

[20] F. J. Martin-Vega, Y. Liu, G. Gomez, M. C. Aguayo-Torres and M. Elkashlan, "Modeling and analysis of NOMA enabled CRAN with cluster point process," IEEE Global Communications Conference, Singapore, 2017, pp. 1-6.

[21] W. Dinkelbach, "On nonlinear fractional programming," *Management Science*, vol. 13, no. 7, pp. 492-498, Mar. 1967.

[22] E. Candes, M.Wakin, and S. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *J. Fourier Anal. Appl.*, vol. 14, no. 5, pp. 877-905, 2008.

[23] S. Boyd and L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004.

[24] Q. Zhang, Q. Li and J. Qin, "Robust beamforming for non-orthogonal multiple access systems in MISO channels," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10231-10236, Dec. 2016.

[25] "Evolved universal terrestrial radio access (E-UTRA); physical channels and modulation (release 12)," document 3GPP Std. 36.211, 2014.

[26] Evolved universal terrestrial radio access (E-UTRA); radio frequency (RF) requirements for LTE pico node B (release 12), document 3GPP Std. 36.931, 2014.