

User Association in Cloud RANs with Massive MIMO

Saeedeh Parsaeefard*, Vikas Jumba[†], Atoosa Dalili Shoaee[‡], Mahsa Derakhshani[‡], Tho Le-Ngoc[‡]

* Iran Telecommunication Research Center, Tehran, Iran

[†]Department of Electrical & Computer Engineering, McGill University, Montreal, QC, Canada

[‡]Wolfson School of Mechanical, Electrical & Manufacturing Engineering, Loughborough University, UK

Email: saeideh.parsaeifard@mcgill.ca; vikas.jumba@mail.mcgill.ca; atoosa.dalilishoaee@mail.mcgill.ca ; m.derakhshani@lboro.ac.uk; tho.le-ngoc@mcgill.ca

Abstract—This paper studies a resource allocation problem where a set of users within a specific region is served by cloud radio access network (C-RAN) structure consisting of a set of base-band units (BBUs) connected to a set of radio remote heads (RRHs) equipped with a large number of antennas via limited capacity front-haul links. User association to each RRH, BBU and front-haul link is essential to achieve high rates for cell-edge users under network limitations. We introduce two types of optimization variables to formulate this resource allocation problem: (i) C-RAN user association factor (UAF) including RRH, BBU and front-haul allocation for each user and (ii) power allocation vector. The formulated optimization problem is non-convex with high computational complexity. An efficient two-level iterative approach is proposed. The higher level consists of two steps where, in each step, one of these two optimization variables is fixed to derive the other. At the lower level, by applying different transformations and convexification techniques, the optimization problem in each step is broken down into a sequence of geometric programming (GP) problems to be solved by the successive convex approximation (SCA). Simulation results reveal the effectiveness of the proposed approach to increase the total throughput of network, specifically for cell-edge users. It outperforms the traditional user association approach, in which, each user is first assigned to the RRH with the largest average value of signal strength, and then, based on this fixed user association, front-haul link association and power allocation are optimized.

Index Terms—Complementary geometric programming, cloud radio access network (C-RAN), fifth-generation (5G) wireless networks, successive convex approximation, massive MIMO.

I. INTRODUCTION

A. Motivations

Emerging fifth-generation (5G) cellular networks have been targeted to provide flexible and software defined based structure to serve diverse applications with different quality-of-service (QoS) requirements. To add flexibility in existing wireless networks, one of the proposed concepts is to change the cell-based access structure of traditional wireless networks with the advent of cloud radio access network (C-RAN) technology [1], in which RF and base-band modules of traditional base stations in one specific region are separated into base band units (BBUs) and radio remote heads (RRHs) [2].

Two major objectives of cloud RAN are to increase the spectral efficiency and to improve the throughput of cell-edge users. Deploying a large number of antennas at the

RRHs, referred to as massive multiple input multiple output (massive MIMO), can bring higher spectrum efficiency to 5G [3]. This scales up the complexity gain of traditional MIMOs [1], [4] and reduces the interference among users of all access points [4]. However, due to interference among users in different coverage regions of RRHs [5], collaborative user scheduling among RRHs is essential to increase the throughput of cell-edge users. Additionally, in such a network, the user association to access points is challenging since the general concept of cell is not applicable to C-RAN, and the RRHs have highly overlapped coverage areas. Furthermore, the limited front-haul link capacity between RRHs and BBUs further complicates the association control as the traditional user association to each RRH based on the signal strength is inadequate.

To address these challenges, in this work, in addition to radio resources (e.g., power), cloud resources (i.e., front haul links and BBUs) are assigned to each user in a joint and centralized manner to maximize the total network throughput under C-RAN limitations. The proposed setup consists of down-link transmission for users in a specific region served by RRHs equipped with massive MIMO. In the optimization formulation, C-RAN-UAF includes BBUs, front-haul link and RRH allocation for each user while power allocation vector determines the allocated power of each RRH to its own users. In this setup, we assume that each user should be connected to one RRH and BBU. We introduce the new constraints to satisfy these practical limitations in the proposed resource allocation problem. Furthermore, the throughput of each user is modified to capture the effects of association of users in BBU, RRHs and front-haul over C-RAN.

Due to user-association constraint and interference among users from different RRHs, the proposed optimization problem is non-convex, suffering from high computational complexity [6]. We apply the frameworks of complementary geometric programming (CGP) and the successive convex approximation (SCA) [7], [8], [9], [10] to develop an efficient iterative algorithm with two-steps to solve the proposed problem. For a given power allocation, Step 1 derives the optimal C-RAN-UAF solution, and subsequently, from obtained C-RAN-UAF, Step 2 derives the optimal power allocation for each user. Even the simplified problem of each step becomes a non-convex optimization problem. We will demonstrate how by applying

various transformation and convexification techniques, such as variable relaxation, DC programming, and CGP, the highly non-convex problem can be transformed into the equivalent lower-bound geometric programming (GP) problems [9], which can be solved via efficient on-line available software, e.g., CVX [11].

To verify the effects of considering C-RAN-UAF, we also propose the resource allocation based on traditional RRH allocation in which each user is assigned to the RRH with the highest value of received signal-to-interference-plus-noise ratio (SINR), and subsequently, front-haul, BBU, and power allocation parameters are derived for the users of each RRH. The latter problem is still non-convex and we propose to apply the CGP again to reach the solution. It should be noted that even though the user association in C-RAN-UAF algorithm is dependent on the channel conditions as in the traditional scheme, the key difference is that the proposed algorithm takes into account the channel conditions of all users in the networks as well as many system requirements as constraints, while in the traditional scheme each user individually makes the association decision only based on its own channel condition. In other words, the traditional scheme is oblivious to the network status, while we are proposing a network-aware joint user association and power allocation algorithm. Simulation results demonstrate that the proposed approach can significantly outperform the traditional approach thanks to a more interference-aware resource allocation via C-RAN-UAF. We also evaluate the computational complexity and the convergence of the proposed algorithm.

B. Related Works

Our work in this paper lies along the intersection of two research contexts in resource allocation problems: 1) load balancing in multi-cell traditional wireless networks, and 2) resource allocation in C-RAN.

Due to users' mobility in cellular wireless networks, user association in multi-cell and/or multi-tier network and load balancing among different access points are important to enhance the network performance. Load balancing is intrinsic in the user association to access points. Traditionally, a user associates to the BS with the maximum received SINR [12]. Such schemes are, in general, sub-optimal in practice especially in the case of non-uniform user distribution or the case of a large number of cell-edge users. There exists a large body of research conducted in resource allocation for multi-cell traditional wireless networks, e.g., in [13], [14], [15], [16], [17], [18], [19], [20]. In this type of works, the BS assignment algorithm is separated from the sub-carrier allocation, while joint sub-carrier and power allocation is applied for multi-cell scenario. In contrast, [21] has reported that the total throughput of network is improved, specifically, the coverage of network at the cell-edge can be improved up to 70% via joint assignment of the BS, sub-carrier and power. In this work, we follow the definition of user association from [21] for the case that the edge of network is equipped with the C-RAN. In this work, C-RAN UAF consists of RRH, BBU and front-haul allocation, while in [21], UAF just consists of BS and

sub-carrier allocation. Therefore, the optimization problem of this paper involves more computational complexity.

The potential of increasing the spectral efficiency of C-RAN via massive MIMO RRHs is reported in recent works. For instance, in [22], two types of data transfer methods for cloud massive MIMO are proposed with the objective to select the best transmission strategy depending on the beam-forming technique, the number of concurrently receiving users and the number of used antennas for transmission. In [23], [24], [25], the beam-forming coordinated interference cancellation and user association are investigated for traditional MIMO based C-RAN. In contrast this paper utilizes the large-scale effect of massive MIMO in C-RAN to simplify the rate formulation in [5], in which the rate of each RRH is related to the numbers of users associated to that RRH and its own number of users. Consequently, the relationship between the total C-RAN rate and considered UAF is highly complex. Via different steps and approximations as well as convexification techniques, we show how the throughput can be simplified.

C. Organization of Paper

The rest of this paper is organized as follows. In Section II, the network setup and problem formulations are introduced. Section III presents the proposed iterative algorithms to solve the formulated optimization problem. Section IV demonstrates the simulation results and their detailed computational complexity analysis, followed by concluding remarks in Section V.

II. NETWORK SETUP AND PROBLEM FORMULATION

We consider a down-link transmission in 5G wireless network with a cloud radio access network (C-RAN) architecture as illustrated in Fig. II where a specific region is covered by a set of $\mathcal{J} = \{1, \dots, J\}$ RRHs. Each RRH is equipped with a large number of antennas, i.e., $M_j \gg 1$ and connected to the C-RAN consisting of $\mathcal{B} = \{1, \dots, B\}$ BBUs. The C-RAN is responsible to process the baseband signals for all RRHs in this region. A limited capacity front-haul link connects RRHs to the BBUs, serving a set of single-antenna users denoted by $\mathcal{N} = \{1, \dots, N\}$ in this specific region.

Define $\alpha_{j,n}$ as a C-RAN user-association factor (C-RAN-UAF) of user $n \in \mathcal{N}$ where

$$\alpha_{j,n} = \begin{cases} 1, & \text{if } n \text{ is allocated to RRH } j, \\ 0, & \text{otherwise.} \end{cases}$$

The transmit power limitation of each RRH j is denoted by P_j^{\max} , and due to this physical-layer limitation, we have

$$C_1 : U_j P_j \leq P_j^{\max}, \quad \forall j \in \mathcal{J},$$

where $U_j = \sum_{n \in \mathcal{N}} \alpha_{j,n}$ is the total number of users associated to the RRH j and P_j is the transmitted power of RRH j . For this setup, we assume that each user n can only associate to one RRH. This implementation limitation can be mathematically represented as

$$C_2 : \sum_{j \in \mathcal{J}} \alpha_{j,n} \leq 1, \quad \forall n \in \mathcal{N}.$$

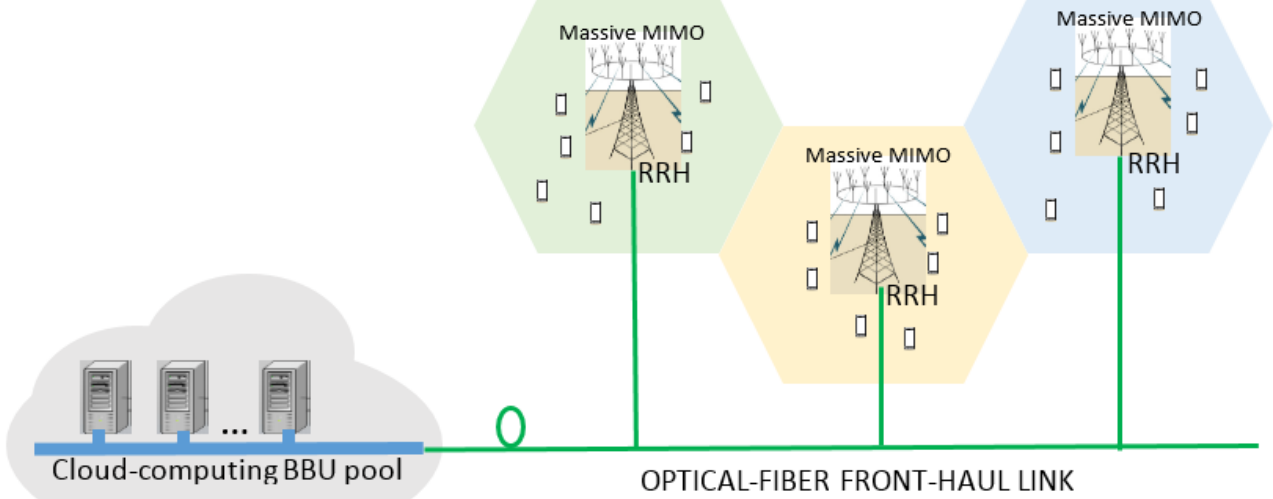


Fig. 1. C-RAN architecture with cloud-computing BBU pool and massive-MIMO RRHs.

Each RRH is connected to C-RAN via limited capacity front-haul link where f_n^b represents the front-haul link association factor between the BBUs' cloud and for users such that

$$f_n^b = \begin{cases} 1, & \text{if user } n \text{ is supported by BBU } b, \\ 0, & \text{otherwise.} \end{cases}$$

Let $g_{j,n}$ be the channel gain of user n to RRH j , and the number of simultaneously served users by a RRH j , be much smaller than the number of transmit antennas M_j . Under this condition, according to [26], the achievable rate of user $n \in \mathcal{N}$ associated to RRH $j \in \mathcal{J}$ can be closely approximated as

$$R_{j,n}(\mathbf{P}, \boldsymbol{\alpha}, \mathbf{F}) = F_n \alpha_{j,n} \log\left(1 + \left(\frac{M_j - U_j + 1}{U_j} \frac{P_j g_{j,n}}{1 + \sum_{j' \neq j} P_{j'} g_{j',n}}\right)\right), \quad (1)$$

where

$$C_3 : F_n = \sum_{b \in \mathcal{B}} f_n^b, \quad \forall n \in \mathcal{N},$$

and, \mathbf{P} , $\boldsymbol{\alpha}$ and \mathbf{F} are the vectors of all P_j , $\alpha_{j,n}$ and f_n^b , respectively, for all $n \in \mathcal{N}$, $b \in \mathcal{B}$ and $j \in \mathcal{J}$.

Note that in (1), we modify the throughput based on the features of C-RAN to include the effect of front-haul link parameter association in C-RAN. Therefore, without allocating both BBUs and RRHs, i.e., when $f_n^b = 0$, the throughput of user is zero.

We assume that each BBU b can handle the maximum load allocation denoted by U_b^{\max} where the allocated load to each BBU is an increasing function of number of assigned users to that BBU [27]. To mathematically represent the load allocation constraint of each BBU, we consider the following

linear function for each BBU load

$$C_4 : \sum_{n \in \mathcal{N}} w_n^b f_n^b \leq U_b^{\max}, \quad \forall b \in \mathcal{B},$$

where w_n^b is the load balancing factor related to BBU b for user n which is a fixed integer number assigned by the network to control the traffic and load of each BBU and user¹. To control the load of C-RAN, we assume that user n is supported by one BBU, and the user n is assigned to BBU if and only if at least one RRH is allocated to that user. We mutually represent these two practical implementation issues as one constraint as

$$C_5 : F_n = \sum_{j \in \mathcal{J}} \alpha_{j,n}, \quad \forall n \in \mathcal{N}.$$

Note that from C_2 , F_n cannot be greater than 1, e.g., if the user n is allocated to one RRH, it is assigned to C-RAN. Otherwise, F_n is equal to 0, which controls the load of C-RAN. The front-haul link between each RRH and the C-RAN has a limited capacity represented by

$$C_6 : \sum_{n \in \mathcal{N}} f_n^b \alpha_{j,n} \leq T_j^b, \quad \forall b \in \mathcal{B},$$

where T_j^b is the transmit front-haul link limitation between RRH j and BBU b . In order to efficiently utilize the power and avoid the situation of power wastage when no user is allocated to any RRH, we consider the following constraint where at least one user is allocated to each RRH.

$$C_7 : U_j \geq 1, \quad \forall j \in \mathcal{J},$$

¹It should be noted that the load of each BBU may not be linear function. For simplify, in this paper, we assume the linear function of load of each BBU. However, with generating the random number w_n^b in simulation, the non-linearity of load function can be covered.

TABLE I
LIST OF KEY NOTATIONS

Notations	Description
\mathcal{J}	Set of RRHs
M_j	Number of antennas of RRH j
\mathcal{B}	Set of BBUs
\mathcal{N}	Set of users
$\alpha_{j,n}$	A C-RAN user-association factor (C-RAN-UAF) of user n at RRH j
P_j^{\max}	transmit power limitation of each RRH j
U_j	Total number of users associated to the RRH j
P_j	Transmitted power of RRH j
f_n^b	Front-haul link association factor between BBU b and user n
$g_{j,n}$	Channel gain of user n to RRH j
U_b^{\max}	Maximum load allocation can be handled by BBU b
w_n^b	Load balancing factor related to BBU b for user n
T_j^b	Transmit front-haul link limitation between RRH j and BBU b

Consequently, based on all of the above network limitations, the resource allocation problem with the objective to maximize the total throughput can be written as the following optimization problem

$$\begin{aligned} \max_{\mathbf{P}, \boldsymbol{\alpha}, \mathbf{F}} \quad & \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} R_{j,n}(\mathbf{P}, \boldsymbol{\alpha}, \mathbf{F}), \\ \text{subject to:} \quad & \mathbf{C}_1 - \mathbf{C}_7. \end{aligned} \quad (2)$$

The formulated problem (2) has non-convex and combinatorial structure with high computational complexity. To overcome this issue, we first relax the integer variables to the continuous ones and then, we resort to SCA, CGP and DC-approximation to transform the non-convex optimization problem into a GP structured problem. Afterwards, we will solve the problem via a two-level iterative algorithm. In the following section, we briefly study the mathematical background and then, propose our algorithm to solve this problem efficiently.

III. TWO-LEVEL ITERATIVE ALGORITHM FOR JOINT USER ASSOCIATION AND POWER ALLOCATION

To tackle the computational complexity of (2), we follow an iterative approach to find C-RAN UAF, i.e., $\boldsymbol{\alpha}$ and \mathbf{F} , and power transmission of each RRH, i.e., \mathbf{P} [28]. This iterative approach contains two layers where at higher-layer two steps exist: 1) the user-association problem is solved for C-RAN UAF vectors as the optimization variables for a given (fixed) power allocation, and 2) the power-allocation problem is solved to get optimal power from the obtained C-RAN UAF vectors in 1). We run the whole algorithm iteratively until we reach to an optimal C-RAN UAF and power allocation. This

sequential resource allocation can further be explained as

$$\underbrace{\boldsymbol{\alpha}(0), \mathbf{F}(0) \rightarrow \mathbf{P}(0) \rightarrow \dots \boldsymbol{\alpha}^*(t), \mathbf{F}^*(t) \rightarrow \mathbf{P}^*(t)}_{\text{Initialization}} \xrightarrow{\text{Iteration } t} \underbrace{\boldsymbol{\alpha}^*, \mathbf{F}^* \rightarrow \mathbf{P}^*}_{\text{Optimal solution}},$$

where $t > 0$ is the iteration index. Also, $\boldsymbol{\alpha}^*(t)$, $\mathbf{F}^*(t)$ and $\mathbf{P}^*(t)$ are optimal values obtained at the iteration t from convex transformation of related optimization problems in each step. The iterative procedure is stopped when

$$\begin{aligned} \|\boldsymbol{\alpha}^*(t) - \boldsymbol{\alpha}^*(t-1)\| &\leq \varepsilon_1, \\ \|\mathbf{F}^*(t) - \mathbf{F}^*(t-1)\| &\leq \varepsilon_2, \text{ and,} \\ \|\mathbf{P}^*(t) - \mathbf{P}^*(t-1)\| &\leq \varepsilon_3, \end{aligned} \quad (3)$$

where $0 < \varepsilon_1 \ll 1$, $0 < \varepsilon_2 \ll 1$ and $0 < \varepsilon_3 \ll 1$.

Notably, both the C-RAN UAF and power allocation problems are still non-convex and suffer from high computational complexity. To solve them efficiently, at the lower-layer of proposed algorithm, we apply complementary geometric programming (CGP) to each step [9] where via different transformations and convexification approaches, the sequence of lower bound GP based approximations of the problem is solved (See Appendix A). Therefore, within each step, another iterative algorithm is applied to convert the non-convex optimization problem to its lower-bound GP based problem in each iteration which has been called lower-layer of the algorithm in this paper. In fact, the outer iterative algorithm is considered as being in the higher level, while the inner iterative algorithms, used to solve the sub-problems 1 and 2, belong to the lower level. Before explaining the details of the user-association and power allocation problem and their solutions, we first review the mathematical preliminaries required to solve the problems in each step in the lower-layer.

A. C-RAN UAF Allocation Algorithm

This problem finds the optimal values of $\boldsymbol{\alpha}(t)$ and $\mathbf{F}(t)$ at each iteration t for users with fixed values of $\mathbf{P}(t)$. In this case, assuming a high signal-to-interference-plus-noise ratio (SINR) scenario, we have

$$\tilde{R}_{j,n}(\boldsymbol{\alpha}, \mathbf{F}) \approx F_n \alpha_{j,n} \log \left(\frac{M_j - U_j + 1}{U_j} \gamma_{j,n}(t) \right),$$

in which $\gamma_{j,n}(t) = \frac{P_j(t)g_{j,n}}{1 + \sum_{j' \neq j} P_{j'}(t)g_{j',n}}$ is the SINR of user n at RRH $j \in \mathcal{J}$ and it has a fixed value for this step. Consequently, the optimization problem related to this step is

$$\begin{aligned} \max_{\boldsymbol{\alpha}, \mathbf{F}} \quad & \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} \tilde{R}_{j,n}(\boldsymbol{\alpha}, \mathbf{F}), \\ \text{subject to:} \quad & \mathbf{C}_1 - \mathbf{C}_7. \end{aligned} \quad (4)$$

In (4), the only optimization variables are $\boldsymbol{\alpha}$ and \mathbf{F} , and therefore, (4) has less computational complexity compared to (2). However, it still suffers from relatively high complexity due to the integer optimization variable $\boldsymbol{\alpha}$ and the non-convex objective function (4). To overcome the computational complexity of (4), we first relax the C-RAN UAF variables as

$\alpha_{j,n} \in [0, 1]$ and $f_n^b \in [0, 1]$. Also, based on the following proposition, we convert (4) into the GP formulation.

Proposition 1: The DC approximation of $\tilde{R}_{j,n}(\alpha, \mathbf{F})$ at iteration t_1 is

$$\tilde{R}_{j,n}(\alpha, \mathbf{F}) \approx F_n(t_1) \alpha_{j,n}(t_1) [\log(M_j \gamma_{j,n}(t_1)) - \log(U_j(t_1 - 1) - \beta_{j,n'}(t_1) + \beta_{j,n'}(t_1 - 1))], \quad (5)$$

where

$$\beta_{j,n'}(t_1) = \sum_{n' \in \mathcal{N}} \frac{\alpha_{j,n'}(t_1)}{\sum_{n' \in \mathcal{N}} \alpha_{j,n'}(t_1 - 1)},$$

and

$$\beta_{j,n'}(t_1 - 1) = \sum_{n' \in \mathcal{N}} \frac{\alpha_{j,n'}(t_1 - 1)}{\sum_{n' \in \mathcal{N}} \alpha_{j,n'}(t_1 - 1)}.$$

Proof. See Appendix B. \square

Based on (5), we can derive the GP counterpart of (4) for each iteration of Algorithm 1 proposed in Table III-B. To obtain a standard GP formulation, the equality constraints in C3 and C7 should involve only monomial functions. In the following, we first relax the variables and then apply iterative AGMA technique to have the monomial approximation for C3 and C7. Also, we show how we can convert the objective function of (5) into the standard form of GP.

Proposition 2: Consider a positive auxiliary variable x_0 and a very large constant value $\Lambda_1 \gg 1$. The GP based reformulation of (4) for each iteration t_1 is

$$\begin{aligned} \mathcal{O}_1(\alpha, \mathbf{F}) : \min_{\alpha} & x_0(t_1), \\ \text{subject to :} & C_{01}, C_1, C_{21}, C_{31}, C_4, C_{51}, C_6, C_7, \end{aligned} \quad (6)$$

where

$$C_{01} : (\Lambda_1 + T_1) \left[\frac{x_0}{a_0(t_1)} \right]^{-a_0(t_1)} T_2 T_3 \leq 1,$$

where

$$T_1 = \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} F_n(t_1) \alpha_{j,n}(t_1) (\log(U_j(t_1 - 1) + \beta_{j,n'}(t_1))),$$

$$T_2 = \prod_{j \in \mathcal{J}, n \in \mathcal{N}} \left[\frac{F_n(t_1) \alpha_{j,n}(t_1) \log(M_j \gamma_{j,n}(t_1))}{a_{j,n}(t_1)} \right]^{-a_{j,n}(t_1)},$$

$$T_3 = \prod_{j \in \mathcal{J}, n \in \mathcal{N}} \left[\frac{F_n(t_1) \alpha_{j,n}(t_1) \beta_{j,n'}(t_1 - 1)}{b_{j,n}(t_1)} \right]^{-b_{j,n}(t_1)}.$$

$$C_{21} : \prod_{n \in \mathcal{N}} \left[\frac{\alpha_{j,n}}{z_{j,n}(t_1)} \right]^{-z_{j,n}(t_1)} \leq 1,$$

$$C_{31} : F_n \prod_{b \in \mathcal{B}} \left[\frac{f_n^b}{q_n^b(t_1)} \right]^{-q_n^b(t_1)} \leq 1,$$

$$C_{51} : F_n \prod_{j \in \mathcal{J}} \left[\frac{\alpha_{j,n}}{v_{j,n}(t_1)} \right]^{-v_{j,n}(t_1)} \leq 1,$$

with the following new variables

$$a_0(t_1) = \frac{x_0(t_1 - 1)}{\tilde{a}_0(t_1)}, \quad (7a)$$

$$a_{j,n}(t_1) = \frac{F_n(t_1 - 1) \alpha_{j,n}(t_1 - 1) \log(M_j \gamma_{j,n}(t_1))}{\tilde{a}_{j,n}(t_1)}, \quad (7b)$$

$$b_{j,n}(t_1) = \frac{F_n(t_1 - 1) \alpha_{j,n}(t_1 - 1) \beta_{j,n'}(t_1 - 1)}{\tilde{b}_{j,n}(t_1)}, \quad (7c)$$

$$q_n^b(t_1) = \frac{f_n^b(t_1 - 1)}{\sum_{b \in \mathcal{B}} f_n^b(t_1 - 1)}, \quad (7d)$$

$$v_{j,n}(t_1) = \frac{\alpha_{j,n}(t_1 - 1)}{\sum_{j \in \mathcal{J}} \alpha_{j,n}(t_1 - 1)}, \quad (7e)$$

$$z_{j,n}(t_1) = \frac{\alpha_{j,n}(t_1 - 1)}{\sum_{n \in \mathcal{N}} \alpha_{j,n}(t_1 - 1)}, \quad (7f)$$

where

$$\tilde{a}_0(t_1) = x_0(t_1 - 1) + \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} F_n(t_1 - 1) \alpha_{j,n}(t_1 - 1) \times (\log(M_j \gamma_{j,n}(t_1)) + \beta_{j,n'}(t_1 - 1)),$$

$$\tilde{a}_{j,n}(t_1) = x_0(t_1 - 1) + \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} F_n(t_1 - 1) \alpha_{j,n}(t_1 - 1) \times (\log(M_j \gamma_{j,n}(t_1)) + \beta_{j,n'}(t_1 - 1)),$$

and

$$\tilde{b}_{j,n}(t_1) = x_0(t_1 - 1) + \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} F_n(t_1 - 1) \alpha_{j,n}(t_1 - 1) \times (\log(M_j \gamma_{j,n}(t_1)) + \beta_{j,n'}(t_1 - 1)).$$

Proof. See Appendix C. \square

Now, the optimization problem $\mathcal{O}_1(\alpha, \mathbf{F})$ can be effectively solved by an available numerical algorithm for GP or on-line available softwares such as CVX [11] which is deployed for this paper to derive the optimal solution of (6).

B. Power Allocation Algorithm

In this section, based on fixed values of $\alpha^*(t)$ and $\mathbf{F}^*(t)$, we propose an algorithm to solve the power allocation problem. Assuming a high SINR scenario, the optimization problem for power allocation is

$$\begin{aligned} \max_{\mathbf{P}} & \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} \tilde{R}_{j,n}(\mathbf{P}), \\ \text{subject to :} & C_1. \end{aligned} \quad (8)$$

Similar to the problem (4), the problem (8) can be translated to its GP counterpart and the iterative algorithm can be applied to solve (8). The GP transformation of (8) is

$$\begin{aligned} \mathcal{O}_2(\mathbf{P}(t)) : \min_{\mathbf{P}} & \prod_{j \in \mathcal{J}, g \in \mathcal{G}, n \in \mathcal{N}_g} \hat{\gamma}_{j,n}(t), \\ \text{subject to :} & \tilde{C}_1(t), \end{aligned}$$

where $\hat{\gamma}_{j,n}(t) = \frac{1}{\gamma_{j,n}(t)} = \frac{1 + \sum_{j' \neq j} P_{j'}(t) g_{j',n}}{M_j P_j(t) g_{j,n}}$ and

$$\tilde{C}_1(t) = U_j(t_1)^* P_j \leq P_j^{\max}, \quad \forall j \in \mathcal{J}^2.$$

²Note that in constraint $\tilde{C}_1(t)$, the value of $U_j(t_1)^*$ is fixed and calculated from the solution of UAF problem.

Algorithm 1: UAF-C-RAN Association and Power Allocation Iterative Algorithm

Initialization: Set $t = t_1 = 0$, t_1^{\max} , t_1^{\max} , $\Lambda_1 = 2000$, $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = \epsilon = 10^{-4}$, $\mathbf{P}(t = 0) = P_j^{\max}/J$,

Repeat: $t = t + 1$

Repeat: $t_1 = t_1 + 1$

Solve $\mathcal{O}_1(\alpha(t_1), \mathbf{F}(t_1))$ for $\alpha(t_1)$ and $\mathbf{F}(t_1)$.

Update CGP variables according to (7)

Until $(\|\alpha(t_1) - \alpha(t_1 - 1)\| < \epsilon$

and $\|\mathbf{F}(t_1) - \mathbf{F}(t_1 - 1)\| < \epsilon$ or $t_1 \geq t_1^{\max}$

Set $\alpha(t) = \alpha(t_1)$, $\mathbf{F}(t) = \mathbf{F}(t_1)$ and reset $t_1 = 0$

Solve $\mathcal{O}_2(\mathbf{P})$ for $\mathbf{P}(t)$.

Until (3) is satisfied or $t \geq t^{\max}$

Now, since $\mathcal{O}_2(\mathbf{P}(t))$ is transformed into the GP, the CVX can be applied to search for $\mathbf{P}^*(t)$.

The overall iterative algorithm to solve C-RAN UAF and power allocation problem is summarized in Algorithm 1. At first, all the required parameters such as iteration parameters and initial values of power are set. Afterwards, in the main loop, two subproblems, C-RAN-UAF and power allocation problems, are solved iteratively until the convergence criteria are met.

C. Sub-optimal Algorithm

As a measure of performance comparison of the proposed algorithm in Algorithm 2, we choose the max SINR approach for user association, where each user is assigned to the RRH with highest average received SINR, similar to traditional wireless networks such as 3G. Mathematically, max SINR based user association problem can be obtained via modifications in (2) as

$$\begin{aligned} \max_{\mathbf{P}, \mathbf{F}} \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} R_{j,n}(\mathbf{P}, \mathbf{F}), \\ \text{subject to: } \mathbf{C}_1, \mathbf{C}_3 - \mathbf{C}_7, \end{aligned} \quad (9)$$

where \mathbf{C}_2 is eliminated since α vector is not a variable here and determined based on the maximum received SINR. Since (9) has two sets of variables, i.e., \mathbf{P} , \mathbf{F} , it has less computational complexity compared to (2) involving three sets of variables. However, it still has a non-convex structure. Again, we propose to decompose (9) into two sub-problems, namely \mathbf{F} allocation and power allocation, respectively.

The \mathbf{F} allocation sub-problem can be written as

$$\begin{aligned} \max_{\mathbf{F}} \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} \tilde{R}_{j,n}(\mathbf{F}), \\ \text{subject to: } \mathbf{C}_3 - \mathbf{C}_6. \end{aligned} \quad (10)$$

To solve this problem, we apply the following proposition.

Proposition 3: Consider t_3 as an iteration index, a positive auxiliary variable $y_0(t_3)$ and a very large constant value $\Lambda_2 \gg 1$. For fixed value of $\Psi = \log\left(\frac{M_j - U_j + 1}{U_j} \gamma_{j,n}\right)$, the GP based reformulation for each iteration t_3 is

$$\begin{aligned} \mathcal{O}_3(\mathbf{F}) : \quad \min_{\mathbf{F}} y_0(t_3), \\ \text{subject to: } \mathbf{C}_{02}, \mathbf{C}_{31}, \tilde{\mathbf{C}}_{52}, \mathbf{C}_6. \end{aligned}$$

Algorithm 2: Max SINR based User Association over C-RAN Algorithm

Initialization: Set $\{t = t_3 = 0, t^{\max} \text{ and } t_3^{\max}\}$, $\Lambda_2 = 2000$, $\varepsilon_2 = \varepsilon_3 = \epsilon = 10^{-4}$, $\mathbf{P}(t = 0) = P_j^{\max}/J$

Repeat: $t = t + 1$,

Set $\alpha(t)$ according to maximum SNR approach.

Repeat: $t_3 = t_3 + 1$

Solve $\mathcal{O}_3(\mathbf{F}(t_3))$ for $\mathbf{F}(t_3)$.

Update CGP variables according to (11) and (12)

Until $\|\mathbf{F}(t_3) - \mathbf{F}(t_3 - 1)\| < \epsilon$ or $t_3 \geq t_3^{\max}$

Set $\mathbf{F}(t) = \mathbf{F}(t_3)$ and reset $t_3 = 0$

Solve $\mathcal{O}_2(\mathbf{P})$ for $\mathbf{P}(t)$.

Until (3) is satisfied or $t \geq t^{\max}$

where

$$\mathbf{C}_{02} : \Lambda_2 \left[\frac{y_0(t_3)}{a_1(t_3)} \right]^{-a_1(t_3)} \prod_{j \in \mathcal{J}, n \in \mathcal{N}} \left[\frac{F_n(t_3) \alpha_{j,n} \Psi}{w_{j,n}(t_3)} \right]^{-w_{j,n}(t_3)} \leq 1.$$

$$\mathbf{C}_{31} : F_n \prod_{b \in \mathcal{B}} \left[\frac{f_n^b}{q_n^b(t_3)} \right]^{-q_n^b(t_3)} \leq 1$$

$$\tilde{\mathbf{C}}_{52} : F_n \leq \sum_{j \in \mathcal{J}} \alpha_{j,n},^3$$

and

$$a_1(t_3) = \quad (11)$$

$$\frac{y_0(t_3 - 1)}{y_0(t_3 - 1) + \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} F_n(t_3 - 1) \alpha_{j,n}(t_3 - 1) \Psi},$$

$$w_{j,n}(t_3) = \quad (12)$$

$$\frac{F_n(t_3 - 1) \alpha_{j,n}(t_3 - 1) \Psi}{y_0(t_3 - 1) + \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} F_n(t_3 - 1) \alpha_{j,n}(t_3 - 1) \Psi}.$$

Proof. See Appendix D. \square

The power allocation problem for this case can be solved similar to $\mathcal{O}_2(\mathbf{P}(t_2))$. The iterative algorithm to solve (9) is summarized in Table III-C, which is similar to Algorithm 1 except that, in the main loop, α is not a variable. In the following section, via simulation results, we will compare performance of the proposed algorithms in Sections III-B and III-C.

IV. SIMULATION RESULTS

This section presents a comprehensive analysis of proposed user association algorithm via simulation results done in Matlab. We consider $N = 12$ users located at random positions inside the region of 4×4 km² served by $J = 4$ RRHs. The values of maximum BBU load, transmission front-haul link limitation and load balancing factor are randomly chosen for each simulation such that $U_b^{\max}, T_j^b \in \{3, 10\}$ and $w_n^b \in \{1, 6\}$. Furthermore, we set total number of BBUs ($B = 4$), number of antennas for each RRH ($M = 200$), maximum power

³Note that in the constraint $\tilde{\mathbf{C}}_{52}$, $\alpha_{j,n}$ is a fixed variable calculated based on the max SINR approach, and hence, it is a valid posynomial function.

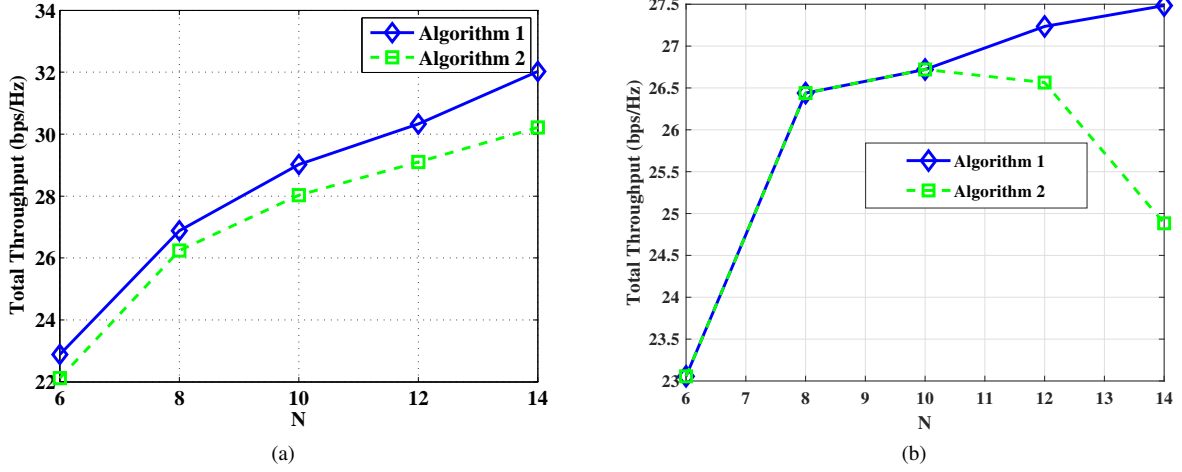


Fig. 2. Throughput versus N with $M = 200$, $J = 4$, and (a) random user locations, (b) fixed user locations

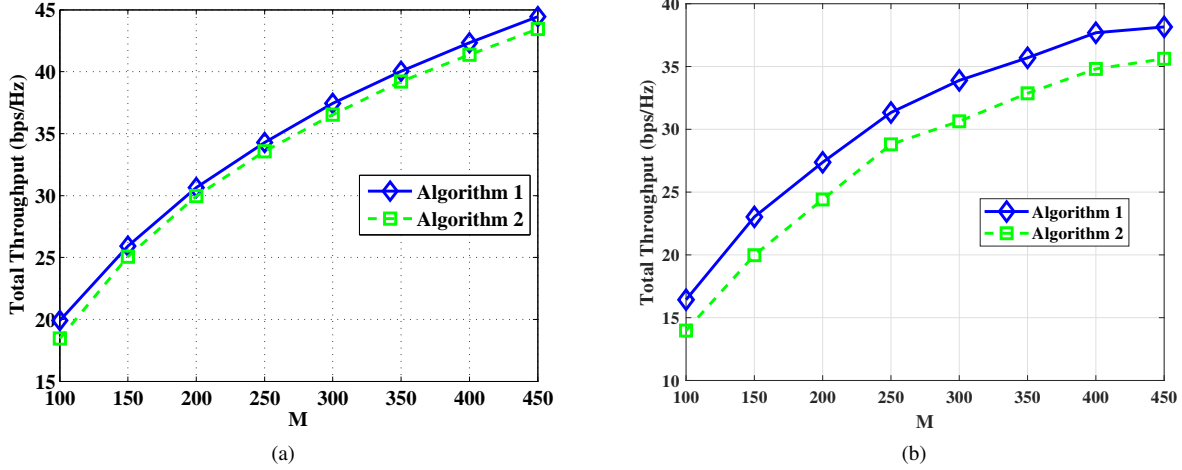


Fig. 3. Throughput versus M with $J = 4$ and (a) $N = 12$, random user locations, (b) $N = 14$, fixed user locations

transmission limit ($P_{\max} = 0$ dB), $\Lambda_1 = \Lambda_2 = 2000$, and $\epsilon_1 = \epsilon_2 = \epsilon_3 = 10^{-3}$ unless otherwise stated. The path-loss from a RRH to j to user n , located at a distance $d_{j,n}$ from RRH, is modeled as $g_{j,n} = \frac{1}{1+(d_{j,n})^4}$ [5]. For scenarios with random location of users, the simulations results are averaged over 100 random realizations, while for scenarios with fixed user locations, the results are averaged over 50 rounds.

To evaluate the performance of Algorithms in Table III-B and III-C, we consider two scenarios: **non-uniformly distributed cell-edge users** where users are only distributed at the edge of cell of each RRHs; and **uniform** where users are uniformly distributed in the entire cell of each RRH.

Fig. 2a illustrates total throughput versus total number of users for Algorithm 1 and 2, where user locations are randomly generated in each simulation round. From Fig. 2a, the total throughput increases with increasing number of users for both approaches as expected from multiuser diversity gain [12]. The higher number of users enhances the opportunity for network to associate more users with

better channel gains to the RRHs, thereby increasing the total throughput of network. The results are also obtained for a scenario with fixed user locations shown in Fig.2b. Unlike Algorithm 1, the performance of Algorithm 2 degrades by increasing the number of users for $N > 10$. The reason is that a larger number of users causes overloaded RRHs which consequently leads to throughput degradation.

Similar to the effect of number of users, the total throughput increases with increasing number of antennas as shown in Figs. 3a and 3b. This is due to the property of multiplexing gain in massive MIMO where the higher number of antennas exploits the multi-path characteristics of wireless channels to transmit parallel streams and multiply the received information via the phenomenon of beam-forming. The results in Fig.3a with random values of U_b^{\max}, T_j^b and also random user locations indicate that Algorithm 2 has a slightly lower performance than Algorithm 1 for the following reasons:

- With Algorithm 2, users are associated to the RRHs based on the maximum SINR. However, due to the front-

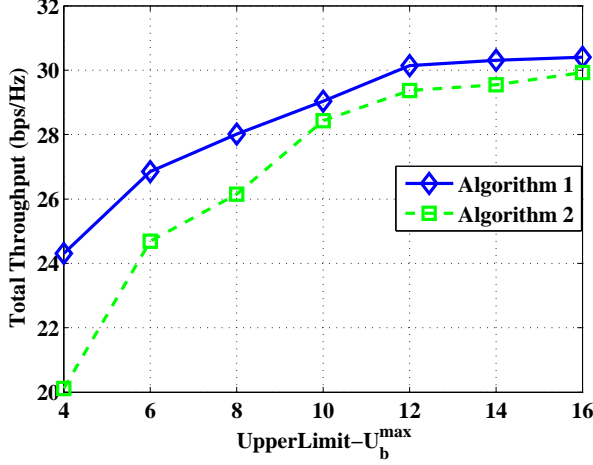


Fig. 4. Rate versus maximum load allocation limit for $N = 12$, $M = 200$, $J = 4$ for non-uniformly distributed cell-edge users

haul link capacity limitation, some of them may not be connected to the network.

- For a large front-haul capacity, based on (1), the achievable rate of an RRH depends on the number of users associated to that RRH. In other words, the term $(M_j - U_j + 1)/U_j$ plays a major role in this condition. In the case of overloaded RRHs, it can be shown that increasing number of users associated to an RRH leads to performance degradation for the RRH at some point.

The superior performance of Algorithm 1 is more pronounced in the case of larger number of users or lower front-haul link capacity as shown in the Fig. 3b with $N = 14$, $U_j = [6]_{(4 \times 4)}$, and fixed user locations. It is noted that the obtained rate is affected by the user-location setting. In fact, the rate depends on how evenly the users are distributed.

Fig. 4 demonstrates the effect of maximum load allocation limit on the total system throughput. Note that the maximum load allocation U_b^{\max} for each BBU in the simulation run is randomly generated in the range from 1 to UpperLimit- U_b^{\max} , shown in the x-axis values in Fig. 4. It is observed that with increasing the range of U_b^{\max} , the total throughput increases. This is because the higher maximum load allocation limit increases the capacity of network to support more users as the load capacity of each BBU is increased. Algorithm 1 outperforms Algorithm 2, indicating the necessity of UAF over limited BBUs capacity in C-RAN. Furthermore, with increasing U_b^{\max} the performance gap between Algorithm 1 and 2 reduces. The reason is that in Algorithm 2, the users are associated to the RRHs based on their received SINRs. Thus, the numbers of users associated to RRHs (or the network load) are not balanced among RRHs. Moreover, there are limitations for each RRH-BBU pair, in constraint C_4 and for each BBU in constraint C_6 . Consequently, for an RRH with a large number of associated users, the performance of Algorithm 2 can be low if its front-haul links with the high capacity are connected to low capacity BBUs. Due to constraint C_6 , some users can be left without any BBU support, i.e. $f_n = 0$. However,

Algorithm 1 jointly allocates the parameters. Therefore, the performance gap between Algorithm 1 and 2 becomes large for small value of U_b^{\max} . However, for the case of larger U_b^{\max} , more users can be supported by each BBU, and Algorithm 2 can offer a better performance.

Figs. 5a and 5b show the cumulative distribution function (cdf) of total throughput for the scenarios of non-uniformly distributed cell-edge users and uniformly distributed users, respectively. It is observed that in both the scenarios, Algorithm 1 offers better performance than Algorithm 2. However, the performance gap between the two algorithms in the scenario of non-uniformly distributed cell-edge users is larger than that in the scenario of uniformly distributed users, and it increases further with increasing number of users as shown in Fig. 5a. For instance, in Fig. 5a, when $N = 20$, Algorithm 2 has lower throughput of 27 bps/Hz as compared to 32 bps/Hz with Algorithm 1 at CDF value of 0.5. This is because when users are located near edges and the network is more congested (higher number of users, e.g., $N = 20$), induced interference to the users of each RRH by other RRHs is increased. In these scenarios, based on the simulations Algorithm 1 can manage the interference more effectively than Algorithm 2 and consequently, Algorithm 1 attains the higher throughput than Algorithm 2.

Fig. 6 shows the total throughput of network versus T_j^b -limit, for $N = 14$ and fixed user locations. In each simulation round, the values of T_j^b are randomly chosen from range of 1 to T_j^b -limit shown on the x-axis in Fig. 6. As observed, with increasing T_j^b , the throughput of both Algorithms 1 and 2 increases since more users can be supported by each front-haul link, and hence the probability that a user is associated to the network due to larger front-haul link capacity also increases. The results also highlight the inefficiency of traditional approach (Algorithm 2) to deal with the limited front-haul capacity as compared to Algorithm 1.

A. Computational Complexity

In this subsection, we analyze the computational complexity of considered UAF and power allocation subproblems in the proposed algorithm in Algorithm 1 via CVX[11]. CVX solves GP problems with the interior point method. The required number of iterations to solve this type of problem is $\log \frac{c/t^0 \rho}{\epsilon}$, where c is the total number of constraints, t^0 is the initial point for approximating the accuracy of interior point method, $0 < \rho < 1$ is the stopping criterion for interior point method and ϵ is used to update the accuracy of interior point method[29].

The number of constraints for UAF and power allocation problem is $c_{\text{UAF}} = 2J + 3N + 2B + 1$ and $c_{\text{PA}} = J$, respectively. The number of computations required for AGMA conversion are $i_{\text{UAF}} = N(J + 1) + B + J$ and $i_{\text{PA}} = 1$, respectively. Note that the subscript UAF and PA denote parameters for UAF and power allocation problems, respectively. Therefore, the total number of computations for the proposed algorithm in Table III-B

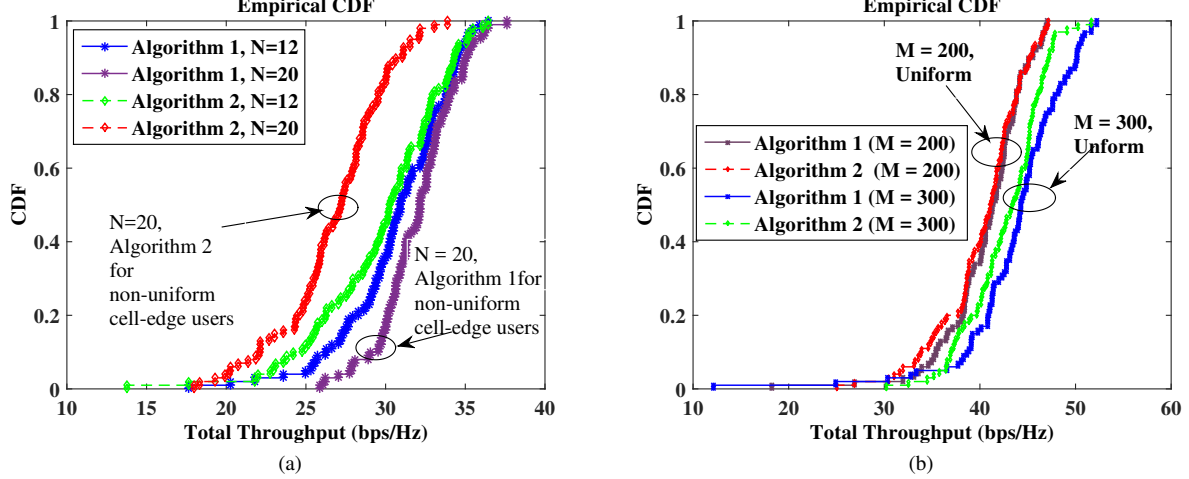


Fig. 5. Throughput Distributions for $J = 4$, different values of M , and (a) non-uniformly distributed cell-edge users with $N = 12, 20$, (b) uniformly distributed users with $N = 12$

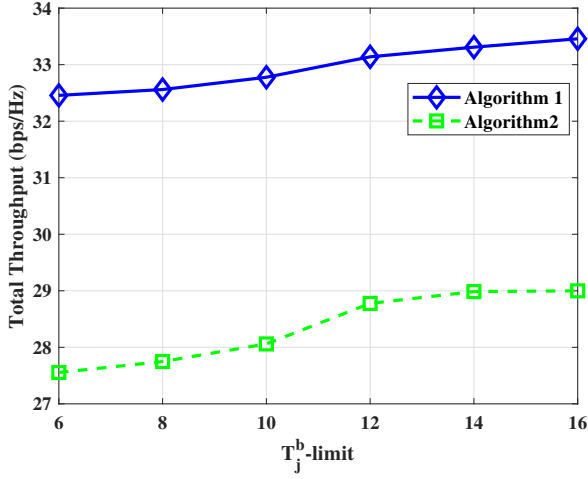


Fig. 6. Rate versus $T_j^b\text{-limit}$ for $N = 14, M = 200, J = 4$, and non-uniformly distributed and fixed-location cell-edge users

$$\text{Computational complexity} = \begin{cases} i_{\text{UAF}} \times \log\left(\frac{C_{\text{UAF}}/t_{\text{UAF}}^0 \rho_{\text{UAF}}}{\epsilon_{\text{UAF}}}\right), \\ i_{\text{PA}} \times \log\left(\frac{C_{\text{PA}}/t_{\text{PA}}^0 \rho_{\text{PA}}}{\epsilon_{\text{PA}}}\right). \end{cases} \quad (13)$$

Based on the above analysis, it can be concluded that the computational complexity of UAF and PA become $O(NJ \log(N+J))$ and $O(\log(J))$, respectively. This implies that the computational complexity of the sub-problems grows only quasi-linearly. The simulation results in Figs. 7a and 7b confirm the analytical analysis of computational complexity according to (13). Similarly, Figs. 8a and 8b illustrate the computational complexity of overall proposed algorithm as $O(1)$ and $O(N)$, with respect to J and N , respectively.

V. CONCLUSION

In this paper, we investigated the user association problem in a cloud RAN, equipped with massive MIMO under the limited capacity of BBUs and front-haul links. In order to reduce the computational complexity, we decomposed the proposed problem into two subproblems: user association factor (UAF) and power allocation. Eventually, a two-level iterative algorithm for user association over C-RAN is developed by using different approximation, CGP and iterative SCA techniques to reduce the computational complexity. Performance of the proposed algorithm is evaluated and compared with the traditional max-SINR based user-association by simulation. Illustrative results show that the proposed algorithm outperforms the traditional approach, especially, for the case of non-uniformly distributed cell-edge users with significant improvement in total achieved system throughput. The simulation results reveal the effectiveness of the proposed approach in dealing with interference between RRHs as well as front-haul capacity limitation.

APPENDIX

A. CGP backgrounds

Geometric programming (GP) is a class of non-linear optimization problems, which can be solved very efficiently via numerical methods [8]. Thus, a significant amount of research has been done in order to convert the resource allocation problems into GP problems, so that it becomes computationally tractable [7], [8], [30], [31], [32].

The standard form of GP is defined as

$$\begin{aligned} \min_{\mathbf{x}} f_0(\mathbf{x}), \quad \text{subject to:} \\ f_i(\mathbf{x}) \leq 1, \quad \forall i = 0, 1, \dots, I, \\ g_j(\mathbf{x}) = 1, \quad \forall j = 0, 1, \dots, J, \end{aligned} \quad (14)$$

where $\mathbf{x} = [x_1, x_2, \dots, x_N]$ is a non-negative optimization variable vector, $g_j(\mathbf{x})$ for all j is a monomial function, i.e., $g_j(\mathbf{x}) = \prod_{n=1}^N c_{jn} x_n^{a_{jn}}$ where $c_{jn} > 0$, $a_{jn} \in \mathbb{R}$, and $f_0(\mathbf{x})$

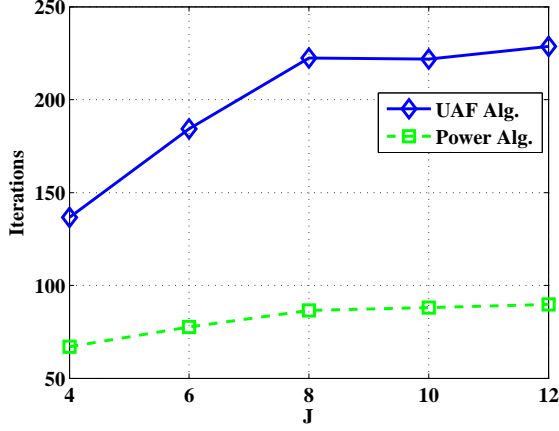
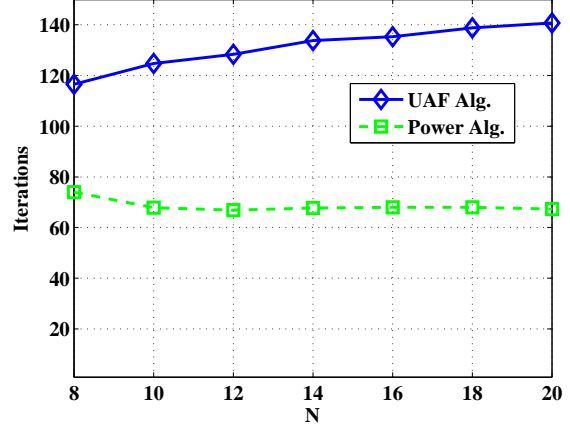
(a) versus J , $N = 12$ (b) versus N , $J = 4$

Fig. 7. Number of CVX Iterations for UAF and Power Allocation Algorithms

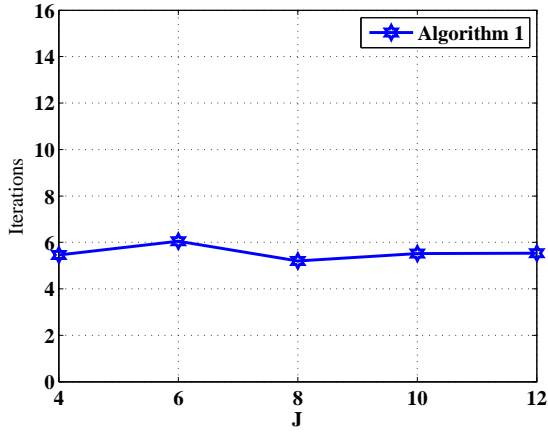
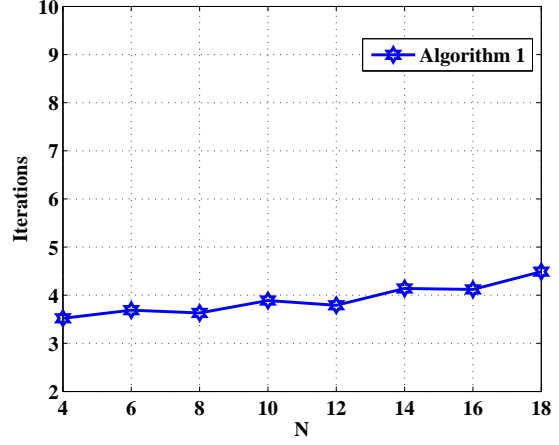
(a) versus J , $N = 12$ (b) versus N , $J = 4$

Fig. 8. Number of iterations for Algorithm 1

and $f_i(\mathbf{x})$ for all i are posynomial functions, i.e., $f_i(\mathbf{x}) = \sum_{k=1}^{K_i} \prod_{n=1}^N c_{ikn} x_n^{a_{ikn}}$. However, in (14), there are a lot of restrictions on the equality and inequality constraints which cannot be met for many practical problems related to the resource allocation of wireless networks such as the optimization problem considered in this paper. For example, in some cases, the equality constraints contain posynomial functions, inequality constraints present lower bound of posynomial function or the posynomial functions contain negative coefficients. Depending on the nature of the optimization problem, these types of problems belong to either one of classes of optimization problems such as generalized GP, signomial programming or complementary geometric programming (CGP). A CGP can be presented as

$$\begin{aligned} & \min_{\mathbf{x}} F_0(\mathbf{x}), \text{ subject to:} \\ & F_i(\mathbf{x}) \leq 1, \quad \forall i = 1, \dots, I, \\ & G_j(\mathbf{x}) = 1, \quad \forall j = 1, \dots, J, \end{aligned} \quad (15)$$

where $F_0(\mathbf{x}) = f_0^+(\mathbf{x}) - f_0^-(\mathbf{x})$, $F_i(\mathbf{x}) = \frac{f_i^+(\mathbf{x})}{f_i^-(\mathbf{x})}$, $i = 1, \dots, I$ and $G_j(\mathbf{x}) = \frac{g_j(\mathbf{x})}{f_j(\mathbf{x})}$ in which $f_0^+(\mathbf{x})$, $f_0^-(\mathbf{x})$, $i = 0, 1, \dots, I$, are posynomial functions, while $g_j(\mathbf{x})$ and $f_j(\mathbf{x})$ are monomial and posynomial functions for $j = 1, \dots, J$, [33], respectively.

One approach to solve (15) is to convert it into a sequence of standard GP problems [9] that can be solved to reach a global solution. In other words, successive convex approximation (SCA) is applied [34], [35], where the non-convex optimization problem is approximated as a convex problem in each iteration. Specifically, arithmetic-geometric mean approximation (AGMA) can be applied to transform the non-posynomial functions to posynomial form, i.e., $F_i(\mathbf{x})$, and $G_j(\mathbf{x})$ to its monomial functions, respectively.

Using AGMA, at the iteration l , the approximated forms of $f_i^-(\mathbf{x}) = \sum_{k=1}^{K_i^-} g_k^{i-}(\mathbf{x})$ and $f_j(\mathbf{x}) = \sum_{k=1}^{K_j} g_k^j(\mathbf{x})$ are

$$\tilde{f}_i^-(\mathbf{x}(l)) = \prod_{k=1}^{K_i^-} \left(\frac{g_k^{i-}(\mathbf{x}(l))}{\alpha_k^{i-}(l)} \right)^{\alpha_k^{i-}(l)}, \quad (16)$$

and

$$\tilde{f}_j(\mathbf{x}(l)) = \prod_{k=1}^{K_j} \left(\frac{g_k^j(\mathbf{x}(l))}{\zeta_k^j(l)} \right)^{\zeta_k^j(l)}, \quad (17)$$

where $\alpha_k^{i-}(l) = \frac{g_k^{i-}(\mathbf{x}(l-1))}{f_i^-(\mathbf{x}(l-1))}$ and $\zeta_k^j(l) = \frac{g_k^j(\mathbf{x}(l-1))}{f_j(\mathbf{x}(l-1))}$. Now, we have $\tilde{F}_i(\mathbf{x}(l)) = \frac{f_i^+(\mathbf{x}(l))}{f_i^-(\mathbf{x}(l))(\mathbf{x}(l))}$ and $\tilde{G}_j(\mathbf{x}(l)) = \frac{g_j(\mathbf{x}(l))}{f_j(\mathbf{x}(l))}$ which are posynomial and monomial functions, respectively [9]. Now, the optimization problem related to each iteration l of (15) is

$$\begin{aligned} \min_{\mathbf{x}(l)} \quad & \Xi + f_0^+(\mathbf{x}(l)) - f_0^-(\mathbf{x}(l)), \\ \text{subject to: } & \tilde{F}_i(\mathbf{x}(l)) \leq 1, \quad \forall i = 1, \dots, I, \\ & \tilde{G}_j(\mathbf{x}(l)) = 1, \quad \forall j = 1, \dots, J, \end{aligned} \quad (18)$$

where $\Xi \gg 1$ is a sufficiently large constant which is added to the objective optimization problem (18) to ensure that the objective function is always positive [9]. However, the objective function of (18) still cannot satisfy the posynomial condition of (14). To reach the GP-based formulation for each iteration, we introduce the auxiliary variable $x_0 > 0$ for a linear objective function and use it to transform (18) into

$$\begin{aligned} \min_{\mathbf{x}_0(t)} \quad & x_0(l), \\ \text{subject to: } & \frac{\Xi + f_0^+(\mathbf{x}(l))}{f_0^-(\mathbf{x}(l)) + x_0} \leq 1, \\ & \tilde{F}_i(\mathbf{x}(t)) \leq 1, \quad \forall i = 0, 1, \dots, I, \\ & \tilde{G}_j(\mathbf{x}(t)) = 1, \quad \forall j = 1, \dots, J, \end{aligned} \quad (19)$$

where $\mathbf{x}_0(t) = [x_0(l), x_n(l), \dots, x_0(l)]$. Similar to $F_i(\mathbf{x})$, term $\frac{\Xi + f_0^+(\mathbf{x}(l))}{f_0^-(\mathbf{x}(l)) + x_0}$ can be converted into posynomial function via AGMA, and finally, the resulting optimization problem has a GP-based structure and can be solved by efficient numerical algorithms [9].

It has been shown that the solution obtained by the iterative algorithm based on the approximation of problem (15) into its GP based approximation has a very close performance to the optimal solution [9].

B. Proof of Proposition 1

From the assumption of $M_j \gg U_j(t_1)$, we will have $\frac{M_j - U_j(t_1) + 1}{U_j(t_1)} \approx \frac{M_j}{U_j(t_1)}$ and the throughput can be approximated to

$$\tilde{R}_{j,n} \approx F_n(t_1) \alpha_{j,n}(t_1) \log \left(\frac{M_j}{U_j(t_1)} \gamma_{j,n}(t) \right). \quad (20)$$

To convexity (20), we apply DC-approximation. Let rewrite (20) as

$$\begin{aligned} \tilde{R}_{j,n}(t_1) \approx \\ F_n(t_1) \alpha_{j,n}(t_1) (\log(M_j \gamma_{j,n}(t)) - \log(U_j(t_1))). \end{aligned} \quad (21)$$

Consider linear approximation of $\log(U_j(t_1))$ as

$$\begin{aligned} \log(U_j(t_1)) \approx \log(U_j(t_1 - 1)) + \\ \nabla \log(U_j(t_1 - 1)) \left(\sum_{n' \in \mathcal{N}} \alpha_{j,n'}(t_1) - \sum_{n' \in \mathcal{N}} \alpha_{j,n'}(t_1 - 1) \right), \end{aligned} \quad (22)$$

where $U_j(t_1) = \sum_{n' \in \mathcal{N}} \alpha_{j,n'}(t_1)$ [36]. Further simplifying (22), we reach to

$$\begin{aligned} \log(U_j(t_1)) \approx \log(U_j(t_1 - 1)) + \\ \sum_{n' \in \mathcal{N}} \frac{\alpha_{j,n'}(t_1)}{\sum_{n' \in \mathcal{N}} \alpha_{j,n'}(t_1 - 1)} - \sum_{n' \in \mathcal{N}} \frac{\alpha_{j,n'}(t_1 - 1)}{\sum_{n' \in \mathcal{N}} \alpha_{j,n'}(t_1 - 1)}. \end{aligned} \quad (23)$$

Substituting (23) into (21), we will have (5).

C. Proof of Proposition 2

In order to solve (23) efficiently, we transform (23) into its equivalent GP formulation. In this context, first, the max based objective function is transformed to a minimization function as shown below

$$\begin{aligned} \min_{\alpha, \mathbf{F}} \quad & \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} -\tilde{R}_{j,n}(\alpha, \mathbf{F}), \\ \text{subject to: } & C_1 - C_7. \end{aligned} \quad (24)$$

Afterward, to obtain a standard GP formulation, the objective function in (24) is transformed to a positive term. In this context, we consider a sufficiently large constant term Λ_1 and positive auxiliary variable $x_0(t_1)$ under the constraint

$$\Lambda_1 - \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} \tilde{R}_{j,n}(t_1) \leq x_0(t_1).$$

Equivalently, after substituting the value of $\tilde{R}_{j,n}(t_1)$ from (5), we obtain the following constraint C_{00} :

$$\begin{aligned} \Lambda_1 + \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} F_n(t_1) \alpha_{j,n}(t_1) (\log(U_j(t_1 - 1)) + \beta_{j,n'}(t_1)) \\ x_0(t_1) + \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} F_n(t_1) \alpha_{j,n}(t_1) (\log(M_j \gamma_{j,n}(t)) + \beta_{j,n'}(t_1 - 1)) \leq 1. \end{aligned}$$

To this end, we reach to the following equivalent optimization problem

$$\begin{aligned} \min_{\alpha, \mathbf{F}} \quad & x_0(t_1), \\ \text{subject to: } & C_{00}, C_1 - C_7. \end{aligned} \quad (25)$$

However, since constraints C_{00}, C_2, C_3 and C_5 in (25) do not satisfy the properties of monomials and posynomials in GP formulations, therefore we apply AGMA approximations to these constraints and reach to $\mathcal{O}_1(\alpha, \mathbf{F})$.

D. Proof of Proposition 3

Similar to Appendix C, consider following equivalent minimization problem of (10)

$$\begin{aligned} \max_{\mathbf{F}} \quad & \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} -\tilde{R}_{j,n}(\mathbf{F}), \\ \text{subject to: } & C_3 - C_6. \end{aligned} \quad (26)$$

In order to reach a standard GP formulation, we consider Λ_2 , a positive constant term and $y_0(t_2)$, an auxiliary variable to define the following constraint

$$\Lambda_2 - \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} R_{j,n} \leq y_0(t_2).$$

After substituting the value of $\tilde{R}_{j,n}(\mathbf{F}) = F_n(t_2)\alpha_{j,n}\Psi$ in the above constraint, we obtain

$$C_{002} : \frac{\Lambda_2}{y_0(t_2) + \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} F_n(t_2)\alpha_{j,n}(t_2)\Psi} \leq 1,$$

where $\Psi = \log(\frac{M_j - U_j + 1}{U_j} \gamma_{j,n})$. The equivalent optimization problem becomes

$$\begin{aligned} & \max_{\mathbf{F}} \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} y_0(t_2), \\ & \text{subject to : } C_{002}, C_3 - C_6. \end{aligned} \quad (27)$$

By applying AGMA to (27), we reach to $\mathcal{O}_3(\mathbf{F})$.

REFERENCES

- [1] V. Jungnickel, K. Manolakis, W. Zirwas, B. Panzner, V. Braun, M. Los-sow, M. Sternad, R. Apelfrojd, and T. Svensson, "The role of small cells, coordinated multipoint, and massive MIMO in 5G," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 44–51, May 2014.
- [2] M. Peng, Y. Li, Z. Zhao, and C. Wang, "System architecture and key technologies for 5G heterogeneous cloud radio access networks," *IEEE Network*, vol. 29, no. 2, pp. 6–14, March 2015.
- [3] F. Boccardi, R. Heath, A. Lozano, T. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, February 2014.
- [4] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, February 2013.
- [5] D. Bethanabhotla, O. Bursalioglu, H. Papadopoulos, and G. Caire, "User association and load balancing for cellular massive MIMO," in *Information Theory and Applications Workshop (ITA)*, February 2014, pp. 1–10.
- [6] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–73, February 2008.
- [7] M. Chiang, "Geometric programming for communication systems," *Foundations and Trends in Communications and Information Theory*, vol. 2, no. 1-2, pp. 1–154, 2005.
- [8] M. Chiang, C. W. Tan, D. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640–2651, July 2007.
- [9] G. Xu, "Global optimization of signomial geometric programming problems," *European Journal of Operational Research*, vol. 233, no. 3, pp. 500–510, 2014.
- [10] M. Derakhshani, X. Wang, D. Tweed, T. Le-Ngoc, and A. Leon-Garcia, "AP-STA association control for throughput maximization in virtualized WiFi networks," to appear in *IEEE Access*.
- [11] I. CVX Research, "CVX: Matlab software for disciplined convex programming, version 2.0," <http://cvxr.com/cvx>, Aug.
- [12] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2004.
- [13] S. Kim, J. Kwon, and J. Lee, "Sum-rate maximization for multi-cell OFDMA systems," *IEEE Trans. Vehicular Technol.*, pp. 1–1, October 2014.
- [14] T. Wang and L. Vandendorpe, "Iterative resource allocation for maximizing weighted sum min-rate in downlink cellular OFDMA systems," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 223–234, January 2011.
- [15] M. Fallgren, "An optimization approach to joint cell, channel and power allocation in multicell relay networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2868–2875, June 2012.
- [16] Y. H. Cho, H. Kim, S.-H. Lee, and H. S. Lee, "A QoE-aware proportional fair resource allocation for multi-cell OFDMA networks," *IEEE Commun. Lett.*, vol. 19, no. 1, pp. 82–85, January 2015.
- [17] J. Yu, G. Y. Li, C. Yin, S. Tang, and X. Zhu, "Multi-cell coordinated scheduling and power allocation in downlink LTE-A systems," in *Proc. IEEE Veh. Tech. Conf. (VTC)*. IEEE, September 2014, pp. 1–5.
- [18] N. Forouzan and S. A. Ghorashi, "New algorithm for joint subchannel and power allocation in multi-cell OFDMA-based cognitive radio networks," *Communications, IET*, vol. 8, no. 4, pp. 508–515, March 2014.
- [19] W. Jing, Z. Lu, X. Wen, Z. Hu, and S. Yang, "Flexible resource allocation for joint optimization of energy and spectral efficiency in OFDMA multi-cell networks," *IEEE Commun. Lett.*, vol. 19, no. 3, pp. 451–454, March 2015.
- [20] N. Ksairi, P. Bianchi, P. Ciblat, and W. Hachem, "Resource allocation for downlink cellular OFDMA Systems-Part I: Optimal allocation," *IEEE Trans. Signal Process.*, vol. 58, no. 2, February 2010.
- [21] S. Parsaeefard, R. Dawadi, M. Derakhshani, and T. Le-Ngoc, "Joint user-association and resource-allocation in virtualized wireless networks," *IEEE Access*, vol. 4, pp. 2738–2750, April 2016.
- [22] S. Park, C. byoung Chae, and S. Bahk, "Before/after precoding massive MIMO systems for cloud radio access networks," *Journal of Communications and Networks*, vol. 15, no. 4, pp. 398–406, August 2013.
- [23] Y. Zhou and W. Yu, "Optimized backhaul compression for uplink cloud radio access network," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1295–1307, June 2014.
- [24] S. Luo, R. Zhang, and T. J. Lim, "Downlink and uplink energy minimization through user association and beamforming in C-RAN," *IEEE Transactions on Wireless Communications*, vol. 14, no. 1, pp. 494–508, January 2015.
- [25] X. Huang, G. Xue, R. Yu, and S. Leng, "Joint scheduling and beamforming coordination in cloud radio access networks with QoS guarantees," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 7, pp. 5449 – 5460, July 2015.
- [26] D. Bethanabhotla, O. Y. Bursalioglu, H. C. Papadopoulos, and G. Caire, "User association and load balancing for cellular massive MIMO," in *Information Theory and Applications Workshop (ITA)*, 2014. IEEE, 2014, pp. 1–10.
- [27] V. N. Ha, L. B. Le, and N.-D. Dao, "Cooperative transmission in cloud RAN considering fronthaul capacity and cloud processing constraints," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, April 2014.
- [28] D. T. Ngo, S. Khakurel, and T. Le-Ngoc, "Joint subchannel assignment and power allocation for OFDMA femtocell networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 342–355, January 2014.
- [29] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2009.
- [30] S. Boyd, S. J. Kim, L. Vandenberghe, and A. Hassibi, "A tutorial on geometric programming," *Optim. Eng.*, vol. 8, no. 1, pp. 67– 127, 2007.
- [31] M. Chiang, D. O'Neill, D. Julian, and S. Boyd, "Resource allocation for QoS provisioning in wireless ad hoc networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, vol. 5, November 2001, pp. 2911–2915.
- [32] M. Chiang and A. Sutivong, "Efficient optimization of constrained nonlinear resource allocation," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, vol. 7, December 2003.
- [33] M. Avriel and A. Williams, "Complementary geometric programming," *SIAM Journal on Applied Mathematics*, vol. 19, no. 1, pp. 125–141, 1970.
- [34] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Operations Research*, vol. 26, no. 4, pp. 681– 683, 1978.
- [35] J. Papandriopoulos and J. S. Evans, "SCALE: A low-complexity distributed protocol for spectrum balancing in multiuser DSL networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3711– 3724, August 2009.
- [36] N. Mokari Yamchi, F. Alavi, S. Parsaeefard, and T. Le-Ngoc, "Limited-feedback resource allocation in heterogeneous wireless networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 4, pp. 2509 – 2521, April 2016.



Saeedeh Parsaeefard (S'09-M'14) received the B.Sc. and M.Sc. degrees from Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran, in 2003 and 2006, respectively, and the Ph.D. degree in electrical and computer engineering from Tarbiat Modares University, Tehran, Iran, in 2012. She was a Post-Doctoral Research Fellow with the Telecommunication and Signal Processing Laboratory in the Department of Electrical and Computer Engineering at the McGill University, Canada. From November 2010 to October 2011, she was a Visiting Ph.D. Student with the Department of Electrical Engineering, University of California, Los Angeles, CA, USA. She is currently a faculty member in the Iran Telecommunication Research Center. Her current research interests include the resource management in software-defined networking, internet of things and the fifth generation of wireless networks as well as applications of robust optimization theory and game theory on the resource allocation and management in wireless networks. She received the IEEE Iran Section Women in Engineering (WIE) awards in 2018.



Vikas Jumba received the B.Tech degree in electronics and communications from Guru Nanak Dev University, Amritsar, India, in 2011, and the MEng. degree in electrical and computer engineering, specialized in telecommunications from McGill University, Montreal, Canada, in 2016. Since 2016, he has been with Ciena as a network core switch developer. Before, he worked as an iOS developer from 2011 to 2013 in Tata Consultancy Services. His areas of interests include future networks and wireless virtualization.”



Tho Le-Ngoc (F97) received the B.Eng. degree (with distinction) in electrical engineering and the M.Eng. degree from McGill University, Montreal, QC, Canada, in 1976 and 1978, respectively, and the Ph.D. degree in digital communications in 1983 from the University of Ottawa, Ottawa, ON, Canada. From 1977 to 1982, he was with Spar Aerospace Ltd., Ste. Anne-de-Bellevue, QC, where he was involved in the development and design of satellite communications systems. From 1982 to 1985, he was an Engineering Manager with the Radio Group, Department of Development Engineering, SRTelecom Inc., Montreal, where he developed the new point-to-multipoint demand assigned time-division multiple-access/time-division multiplexing subscriber radio system SR500. From 1985 to 2000, he was a Professor with the Department of Electrical and Computer Engineering, Concordia University, Montreal. Since 2000, he has been with the Department of Electrical and Computer Engineering, McGill University. His research interests include broadband digital communications. Dr. Le-Ngoc is a Fellow of the Engineering Institute of Canada, the Canadian Academy of Engineering, and the Royal Society of Canada. He received the 2004 Canadian Award in Telecommunications Research and the 2005 IEEE Canada Fessenden Award. He holds a Canada Research Chair (Tier I) on Broadband Access Communications.

Atoosa Dalili Shoaee received her B.Sc. degree in information technology engineering from Isfahan University of Technology, Isfahan, Iran, in 2009 and her M.Sc. degree in information technology engineering from Amirkabir University of Technology, Tehran, Iran, in 2012. Since 2014, she has been with the Department of Electrical and Computer Engineering, McGill University, Canada, where she is currently working toward her Ph.D. degree. Her main areas of research interest are media access control techniques, Internet of Things (IoT) and wireless virtualization.



Mahsa Derakhshani (S10-M13) received the B.Sc. and M.Sc. degrees from Sharif University of Technology, Tehran, Iran, in 2006 and 2008, respectively, and the Ph.D. degree from McGill University, Montreal, Canada, in 2013, all in electrical engineering. From 2013 to 2015, she was a Postdoctoral Research Fellow with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, Canada, and a Research Assistant with the Department of Electrical and Computer Engineering, McGill University. From 2015 to 2016, she was an

Honorary NSERC Postdoctoral Fellow with Department of Electrical and Electronic Engineering, Imperial College London. She is currently a Lecturer (Assistant Professor) in Digital Communications with the Wolfson School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University. Her research interests include radio resource management for wireless networks, software-defined wireless networking, applications of convex optimization and game theory for communication systems, and spectrum sensing techniques in cognitive radio networks. Dr. Derakhshani received the John Bonsall Porter Prize, the McGill Engineering Doctoral Award, the Fonds de Recherche du Québec Nature et Technologies (FRQNT) and Natural Sciences and Engineering Research Council of Canada (NSERC) Postdoctoral Fellowships. Currently, she serves as an Associate Editor of the IET Signal Processing Journal.