Resource Scheduling in Non-Orthogonal Multiple Access (NOMA) based Cloud-RAN Systems

Rupesh Rai, *Student Member*, *IEEE*, H. Zhu, *Senior Member*, *IEEE* and Jiangzhou Wang, *Fellow*, *IEEE*
School of Engineering and Digital Arts, University of Kent, Canterbury, CT2 7NT, United Kingdom.
Email: {rr304, H.Zhu, J.Z.Wang}@kent.ac.uk

***Abstract*—Cloud radio access network (C-RAN) is enabling innovation which can address a number of challenges that cellular network operators (CNOs) face while trying to support ever growing end users' needs towards 5th generation (5G) of mobile networks.This paper investigates a scheduling problem to enhance the performance of non-orthogonal multiple access (NOMA) based cloud radio access networks (C-RAN). We consider the downlink of a C-RAN where cloud is responsible for the scheduling policy and formulate the scheduling problem. In this paper non-orthogonal multiple access (NOMA) is adopted and resource scheduling techniques in an unconstrained C-RAN networks are proposed. We design an efficient resource scheduling algorithm to assist the resource scheduling process in NOMA based C-RAN systems. We formulate the problem as binary integer programming (BIP) optimization problem and propose a suboptimal BIP algorithm and a greedy heuristic allocation algorithm which leads to the maximization of throughput in terms of sum rate when compared with the standard techniques. Simulation results shows significant performance improvement with the proposed algorithms.**

I. Introduction

Cloud radio access network (C-RAN) or dynamic distributed antenna systems (DAS) [1-6] is enabling innovation which can address a number of challenges that cellular network operators (CNOs) face while trying to support ever growing end users' needs towards 5th generation (5G) of mobile networks. Therefore the investigation of the resource allocation and scheduling algorithms in C-RAN networks and its coexistence with other state of the art technology has become the sole impetus of many research groups. A C-RAN provides a new cost-effective way to achieve network densification where the conventional base stations (BSs) are replaced by low-power and low-complexity remote radio heads (RRHs) or remote antenna units that are coordinated by a central processor. In C-RAN joint processing operations are performed at the central processor which leads to increased spectral and energy efficiency via centralized resource allocation [7].

Centralizing baseband processing facilitate better coordination across the remote radio heads (RRHs) as the cell site information like channel conditions, user requirements and traffic loads are available across the network. Such information can be used for optimization of radio resource allocation, manage intercell interference and improve coverage. Sharing information can enhance capacity by enabling the implementation of new technologies such as coordinated multipoint (CoMP) and self organizing networks (SON). As cellular network operators (CNOs) need to continually enhance the capacity of their mobile networks, sharing radio resources between multiple CNOs supports higher peak rates.

In cellular wireless networks, the multiple access (MA) technology is one of the important aspects in improving system capacity. In order to enhance spectrum efficiency in wireless networks, NOMA has been proposed in recent work [8,9]. In NOMA, multiple users are multiplexed by superposition coding in the power domain on the transmitter side and employ SIC to seperate multi-user signal on the receiver side. The system-level performance of downlink NOMA and potential issues (i.e. candidate user set selection, power allocation, error propagation for SIC) are investigated in [10,11]. In order to enhance user fairness in cellular downlink, the proportional fair (PF) based scheduling is introduced in NOMA [12]. In [13], the achievable system throughput of the NOMA and SIC with an assumption that all the resources are proportionally distributed with multi-user scheduling is investigated. In [14] novel interference cancellation (IC) is proposed for NOMA systems.

Most of the resource allocation work in C-RAN is based on orthogonal multiple access schemes such as OFDMA. These OFDMA-based C-RAN systems impose a constraint that each resource block in the network is able to serve at most one user at a time. However to the best of author's knowledge this work is the first attempt to investigate the user scheduling and resource allocation schemes in NOMA-based C-RAN systems.

Resource allocation [15-18] is the core task in order to coordinate transmissions. Traditional distributed solutions for resource allocation are sub-optimal since they rely on the local network information at each serving base station. Such distributed networks lack in the global network information particularly from the nearby base stations. Due to the centralized scheme of C-RAN many tasks such as resource allocation can benefit from the centralized global perspective.

Some recent work combines NOMA with different techniques such as coordinated multi-point (CoMP) [19,20] for future throughput improvement. Hybrid beamforming is discussed in [21]. Recent work in [22] proposes a NOMA scheme for wireless downlink C-RAN. To the best of author's knowledge this work is the first attempt to investigate the resource allocation and scheduling in NOMA integrating with C-RAN.

In this paper we propose a NOMA based efficient resource scheduling mechanism to support multiple CNOs in a C-RAN taking into account the number of UEs and RRHs. By using the resource scheduling schemes, a higher throughput can be achieved with NOMA as compared to the standard techniques. We also propose a greedy resource scheduling algorithm to schedule resources in NOMA based C-RAN system for a large number of RRHs.

The rest of the paper is organized as follows: In Section II we describe the system model and problem formulation. In Section III we propose power allocation strategies in NOMA

978-1-5386-1104-3/17/$31.00 ©2017 IEEE

system. In IV two efficient resource scheduling algorithms to solve the optimization problem are discussed. Section V describes the performance of the proposed schemes.Conclusions are drawn in Section VI.

II. System Model and Problem Formulation

In this section we present a system model of C-RAN utilising NOMA, derive throughput of NOMA scheme for C-RAN and formulate the optimization problem.

A. System model of C-RAN utilising NOMA

Consider a downlink of a cloud based radio access network architecture shared between K celluar network operators (CNOs) and each kth CNO, $k = 1, 2, ..., K$ has N_k remote radio heads (RRHs)$\{RRH_{k,1}, RRH_{k,2}........,RRH_{k,N_k}\}$. The distance between RRH_{k,j_k}, $k = 1, 2....K$, $j_k = 1, 2....N_k$ and the pool of BBUs is dnoted by d_{k,j_k}. It is assumed that the RRHs are connected to a pool of Baseband Units (BBUs) in remote data centres via transport networks such as optical transport network and the signalling is perfectly synchronized. We consider one antenna per network point (RRH/user), having the same transmit power operating on the same system bandwidth B and assume perfect channel state information (CSI) at both transmitters and receivers. Non-orthogonal multiple access (NOMA) is used for downlink transmission. The number of resource blocks available in a network is R. The total number of UEs in a network is U_i. In this paper all the RBs are homogeneous. This means that the throughput of UE is same irrespective of the RB which is allocated to them. In our model the system scheduler allocates the RBs to the UEs during transmission time interval (TTI). The UE set $S = \{S_R\}$ represents the scheduling set where $S_R = \{U_1, U_2...., U_i\}$ is the set of UEs served by the RRHs based on NOMA based C-RAN scheme.

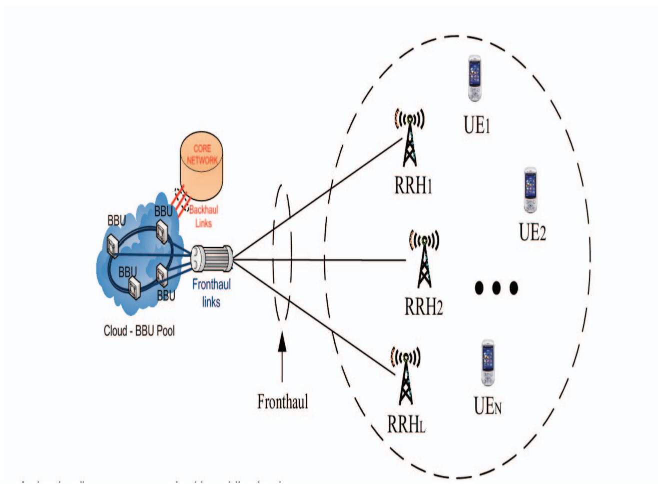

Figure 1. System Model

Let us denote B as the transmission bandwidth of the wireless downlink. The signal transmitted from BBU to RRH_{k,j_k} is x_{k,j_k}. The number of bits that can be transmitted over the RB is determined by the SINR which depends on channel fading parameters, power and interference due to other RRHs. The downlink signal to interference plus noise ratio SINR of a link between UE u and RRH_{k,j_k} $k = 1, 2....K$, $j_k = 1, 2....N_k$ can be expressed as:

$$\gamma_{k,j_k} = \frac{P_{k,j_k}|h_{k,j_k}|^2}{\sum\limits_{c\in C_n} P_{k,j_k}|h_{k,j_k}|^2 + \sigma^2} \tag{1}$$

where P_{k,j_k} is the transmit power of RRH_{k,j_k}, h_{k,j_k} is the channel gain of wireless link. C_n is the set of RRHs that interfere with RRH_{k,j_k}.

The signals for all RRH_{k,j_k} are superimposed to the pool of BBUs as follows:

$$x = \sum_{k=1}^{K}\sum_{j_k=1}^{N_k}\sqrt{P_{k,j_k}}x_{k,j_k} \tag{2}$$

The signal received at RRH_{k,j_k} is given by:

$$y_{k,j_k} = h_{k,j_k}x + n_{k,j_k} \tag{3}$$

where n_{k,j_k} is the complex Gaussian noise at the RRH_{k,j_k} having zero mean and variance σ^2_{k,j_k}.

From (2) and (3) the signal received at RRH_{k,j_k} can be expressed as:

$$y_{k,j_k} = \sum_{k=1}^{K}\sum_{j_k=1}^{N_k} h_{k,j_k}\sqrt{P_{k,j_k}}x_{k,j_k} + n_{k,j_k} \tag{4}$$

The maximum number of bits that can be transmitted at the RRH_{k,j_k} with NOMA can be computed by:

$$R^{NOMA}_{k,jk} = B\log_2\left(1 + \frac{P_{k,j_k}|h_{k,j_k}|^2}{\sum_{l=1}^{j_k-1}P_{k,l}|h_{k,l}|^2 + \sigma^2_k}\right) \tag{5}$$

B. Problem Formulation

The aim of the cellular network operator (CNO) is to maximize its sum weighted data rates.The weights are selected by CNOs according to their scheduling policies. Assuming that UE is connected to RRH_{k,j_k}, the scheduling problem can be formulated as:

$$max\sum_{j=1}^{N_k}\sum_{k}\left[\sum_{U\in U_{k,j_k}}\sum_{r=1}^{R}\hat{w_u}R^{NOMA}_{k,j_k}\beta_{k,j_k}\right] \tag{6}$$

subject to

$$\sum_{c\in C_n}\sum_{u\in U_c}\beta_{r,j_k} + \sum_{u\in U_i}\beta_{r,j_k} \leq 2, \forall_{k,r} \tag{7}$$

where R_k, j_k^{NOMA} is the data rate of NOMA achieved by assigning RB r to UE u, $\hat{w_u}$ is the normalized weight for UE u, $U_{j_k} = \cup_k U_{k,j_k}$ is the set of UEs connected to $RRHN_k$.

III. Power Allocation

In order to ensure the QoS of users and give importance to users who are far from reaching thier requested data rates, a weight is assigned to each user based on the distance between its requested data rate and its actual achieved throughput as follows:

$$w_1 = \frac{(|R_{1,QoS} - R_{1,Obt}|)^{2\alpha}}{\sum_{k\in K,i=1,2}(|R_{i,QoS} - R_{i,Obt}|)^{2\alpha}} \tag{8}$$

Similarly

$$w_2 = \frac{(|R_{2,QoS} - R_{2,Obt}|)^{2\alpha}}{\sum_{k\in K,i=1,2}(|R_{i,QoS} - R_{i,Obt}|)^{2\alpha}} \tag{9}$$

where α is the control parameter used in simulations. $R_{1,QoS}$ and $R_{2,QoS}$ are the data rates required to meet the QoS. $R_{1,Obt}$ and $R_{2,Obt}$ are the actual data rates obtained. The optimal and suboptimal power allocation methods have been proposed and any of them can be applied in the proposed system. As the main target is to schedule the users to enhance the total throughput of the system we consider a simple power allocation method in this paper. The transmission power of the uth user is presented as:

$$P_{k,j_k} = \sqrt{\frac{|h_{j_k,U_N-u+1}|^2}{\sum_{u=1}^{U_N}|h_{j_k,u}|^2}} \tag{10}$$

In (10), the transmission power of the uth user is inversely proportional to the channel gain.The lowest power is allocated to the user closest to RRH and the highest power is allocated to the user farthest from the RRH. In power domain NOMA, user's data multiplexing is performed in the power domain. The sum of assigned power allocation for all users sharing the same frequency block is expressed as:

$$\sum_{i=1}^{M} P_i = P \tag{11}$$

Using (5) the SINR of the high priority user can be expressed as:

$$\gamma_{k,j_k} = \frac{P_{high}|h_{high}|^2}{P_{low}|h_{high}|^2 + I_{N_m} + \sigma^2_{k,j_k}} \tag{12}$$

P_{high} and P_{low} are the power allocation coefficients for high and low priority users respectively. Assuming M=2 in (11), we have the following:

$$P = P_{high} + P_{low} \tag{13}$$

Substituting (13) in (12) we get:

$$\gamma_{k,j_k} = \frac{P_{high}|h_{high}|^2}{(P_k - P_{high})|h_{high}|^2 + I_{N_m} + \sigma^2_{k,j_k}} \tag{14}$$

Considering two users, we get power allocation coefficients of the high priority user as follows:

$$P_{high} = \frac{\gamma_{k,j_k}(|h_{high}|^2 + 1/\rho)}{|h|^2(\gamma_{k,j_k}+1)}P \tag{15}$$

where ρ is transmit SNR given by:

$$\rho = \frac{P}{\sigma^2_{k,j_k}} \tag{16}$$

Substituting (13) in (15) and rearranging we get:

$$P_{low} = \frac{(|h_{high}|^2 - \gamma_{k,j_k}/\rho)}{|h_{high}|^2(\gamma_{k,j_k}+1)}P \tag{17}$$

IV. User Scheduling and Resource Allocation

Various scheduling policies are proposed in wireless networks for achieving maximum spectrum utilization, QoS satisfaction and fairness between UEs. Some of them include channel-aware policies such as Maximum Throughput (MT), Proportional Fair (PF) and Generalized PF (GPF). The scheduling problem in (9) is a BIP problem which is NP hard. The optimal solution is achieved by jointly allocating RBs to users who are subscribed to different CNOs across all the available RRHs.

In this section two low complexity algorithms are proposed. The concept behind the algorithm is to maximize the weighted sum utility. The RB is allocated to the user in RRH who can maximize sum utility. To allocate RB r, the set of CNOs at every RRH is expressed by $1 \times N_k$ vector $s_r = [s_{1,r}, s_{2,r}....., s_{j_k,r}]$ where $s_{j_k,r}$ is index of the CNO at RRH N_k. Vectors $a_r = [a_{1,r}, a_{2,r}, ..., a_{N_k,r}]$ and $b_r = [b_{1,r}, b_{2,r}, ..., b_{N_k,r}]$ denotes the utilities and indices of the UEs respectively where:

$$a_{j_k,r} = max_{u \in U_{s_k,j_k}} \hat{w}_u R^{NOMA}_{k,j_k} \tag{18}$$

$$b_{j_k,r} = argmax_{u \in U_{s_k,j_k}} \hat{w}_u R^{NOMA}_{k,j_k} \tag{19}$$

The optimization problem per RB is the Maximum Weighted Set. For each RB the MWS I_s can be found by using the optimization problem:

$$I_s = max \sum_{j=1}^{N_k} v_{j_k}\beta_{j_k} \tag{20}$$

subject to

$$\sum_{c \in C_n} \beta_c \leq 2, \forall k \tag{21}$$

where β_{j_k} is the binary variable, equal to one if $N_k \in I_S$ and zero otherwise.

Algorithm 1 shows the pseudo code of low complexity algorithm. Λ_{k,j_k} is the number of RBs allocated to CNO k at $RRHN_k$.The MWS of RRH that maximizes the sum weight utility is assigned the RB. Lines 5-8 find the CNO (s_{j_k}) at every RRH. Lines 9-10 finds the index and utility of the user who maximizes the sum utility. In line 12, the algorithm solves the optimization problem. The RB is assigned to user in line 14. The algorithm runs until all RBs have been assigned.

Algorithm 1: Proposed RB allocation Algorithm under NOMA+C-RAN

1. **input** Λ_{k,j_k}, $\bar{\Omega}_{k,j_k}$,$\forall k, j_k$
2. $R_{j_k} = R, \forall j_k$
3. **for** $r = 1 : R$ **do**
4.   **for** $j_k = 1 : N_k$ **do**
5. $v_{j_k,r} = 0$
6. $\Delta\Omega_{k,j_k} = \Lambda_{k,j_k} - \bar{\Omega}_{k,j_k}, \forall k$
7. $\Omega_{k,j_k} = \Lambda_{k,j_k} + \Delta\Omega_{k,j_k}, \forall k$
8. $s_{j_k} = argmax_k \Omega_{k,j_k}, \forall k$
9. $a_{j_k,r} = max_{u \in U_{s_k,j_k}} \hat{w}_u R^{NOMA}_{k,j_k}$
10. $b_{j_k,r} = argmax_{u \in U_{s_k,j_k}} \hat{w}_u R^{NOMA}_{k,j_k}$
11.   **end for**
12.   solve the BIP problem in (13)
13.   assign RB r to UE $b_{j_k,r}, \forall j_k \in I_s$
14. $\Omega_{s_k,j_k} = \Omega_{s_k,j_k} - 1, \forall j_k \in I_s$
15.   update $\bar{\Omega}_{s_k,j_k}, \forall j_k \in I_s$
16. **end for**

However for a large number of RRHs, a low complexity greedy Algorithm 2 is shown. The Algorithm 2 is greedy in the sense that it assigns an RB to the RRH that has the maximum weight v_{j_k}. The first 11 lines of the low complexity greedy Algorithm 2 are similar to Algorithm 1 where CNOs and their users who maximize the sum utility are specified. The RRH index j^*_k that has maximum utility is found in line 14.

This is added to the subset I_s in the next line. The RRH j_k^* and its interfering RRHs indices are removed from the set of RRHs S_{ind}, thereby eliminating interference as the interfering RRHs are not assigned to the same RBs. RB is assigned to $b_{j_k,r}, \forall j_k \in I_s$ in line 18 and $\Omega_{s_k,jk}$ is updated in lines 19-20.

Algorithm 2: Proposed Greedy allocation algorithm under NOMA+C-RAN

1. **input** $\Lambda_{k,j_k}, \bar{\Omega}_{k,j_k}, \forall k, j_k$
2. $R_{j_k} = R, \forall j_k$
3. **for** $r = 1 : R$ **do**
4.   **for** $j_k = 1 : N_k$ **do**
5. $v_{j_k,r} = 0$
6. $\Delta\Omega_{k,j_k} = \Lambda_{k,j_k} - \bar{\Omega}_{k,j_k}, \forall k$
7. $\Omega_{k,j_k} = \Lambda_{k,j_k} + \Delta\Omega_{k,j_k}, \forall k$
8. $s_{j_k} = argmax_k \Omega_{k,j_k}, \forall k$
9. $a_{j_k,r} = max_{u \in U_{s_k,j_k}} \hat{w_u} R_{k,j_k}^{NOMA}$
10. $b_{j_k,r} = argmax_{u \in U_{s_k,j_k}} \hat{w_u} R_{k,j_k}^{NOMA}$
11.   **end for**
12. $S_{ind} = \{1, 2, ..., N_k\}$
13.   **while** $S_{ind} \neq \phi$**do**
14. $j_k^* = argmax_{j_k} v_{j_k,r}$
15. $I_s \leftarrow j_k^*$
16. $S_{ind}^* = S_{ind} \backslash c \in \{C_n \cup j_k^*\}$
17.   **end while**
18.   assign RB r to UE $b_{j_k,r}, \forall j_k \in I_s$
19. $\Omega_{s_k,j_k} = \Omega_{s_k,j_k} - 1, \forall j_k \in I_s$
20.   update $\bar{\Omega}_{s_k,j_k}, \forall j_k \in I_s$
21. **end for**

V. Simulation Results

In this section performance of our proposed algorithms for NOMA in C-RAN is evlauated with system level simulations. The simulations are carried out with MATLAB under different scenerios by varying the number of RRHs and UEs. Results are compared with existing static sharing and heuristic schemes adopted in OFDMA. We consider a downlink cellular network with 19 cells deployed in a hexagonal grid pattern for the simulation. The number of users for each CNO at each cell is assumed to be uniform random variable. UEs are assumed to be uniformly distributed across each cell and have average SINR of 5 to 10 dB. The simulation parameters are listed in Table 1.

Table I
Simulation parameters

Parameters	Values
Cell Layout	Hexagonal 19 cell model
Channel Model	Rayleigh Fading Model
Radius of RRH Coverage	500m
RRH transmitter antenna	1
User terminal receiver antenna	1
CNO Scheduler	PF
Scheduling Interval	1ms
RB Bandwidth	180KHz
Path Loss Model	133.6+35log_{10}(d[km])dB
Throughput Calculation	Based in Shannon's Formula
Average SINR	5-10 dB

The proposed schemes are compared with static sharing and heuristic schemes based on OFDMA in which each CNO at each RRH receives its share of RBs and allocates them to its users according to the CNO scheduling policy. In the proposed schemes the spectrum resources are shared between all UEs connected to the RRH according to the NOMA scheme. Numerical results are presented for the scenerio in which each CNO is fully loaded and operates at its full capacity. This scenerio considers the optimal solution and the results of the BIP and greedy algorithms are compared with static sharing and heuristic schemes.

Fig. 2 compares the average throughput per UE for different average number of UEs per cell. As it can be noted from the figure that increasing the average number of UEs per cell decreases the average number of RBs assigned to each UE, which decreases the average throughput per UE. Each scheme provides its own throughput which is based on the adopted resource allocation scheme. All the results are upper bounded by the optimal scheme and lower bounded by the static sharing scheme. Moreover, the proposed BIP and greedy schemes outperforms the static sharing and heuristic schemes and they are only 9% less than the optimal. The throughput advantage of the proposed BIP and the greedy schemes is due to the fact that UEs in static sharing may access only an RB that is assigned to their RRH. Hence UEs loose the chance to use the underutilized RBs that belong to other RRHs. However in the proposed schemes the RBs for one RRH are utilized by the other RRHs or CNOs.

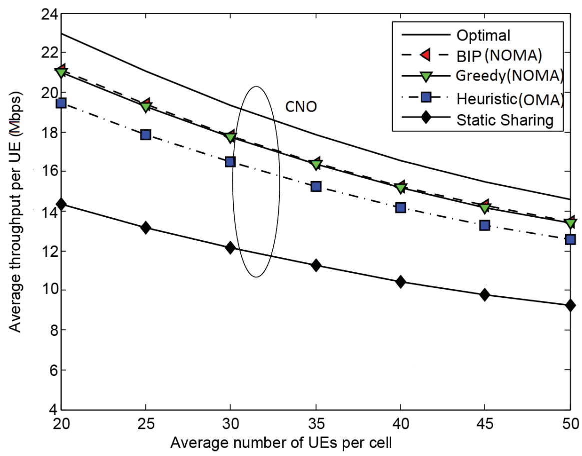

Figure 2. Average Throughput per UE

Fig.3 shows the average aggregate throughput per cell for the different schemes. As the number of UEs increases, the average aggregate throughput increases as a result of the multi CNOs multiplexing gain.

Fig. 4 illustrates the achievable total throughput of the proposed schemes in NOMA based C-RAN system. The total throughput is plotted as a function of number of RRHs which is assumed to vary upto 50 RRHs. It is assumed that the distances between RRHs and the centre stations are in the range of 50m-9km with channel gain from 20 to 0 dB. The total transmission bandwidth is 10 MHz. It can be observed in Fig.4 that the total throughput of the proposed schemes is higher than the static sharing and heuristic schemes. It is

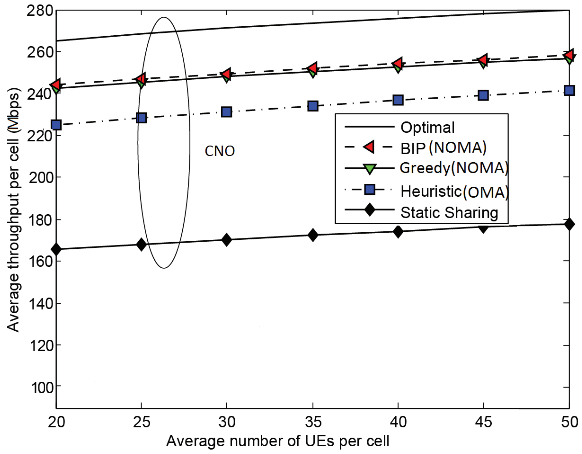

Figure 3. Average Throughput per Cell

noted that with the increase in the number of RRHs in the proposed schemes, there is an increase in the throughput upto certain level, referred to as maximum achievable throughput. Further increase in the number of RRHs will maintain the same throughput.

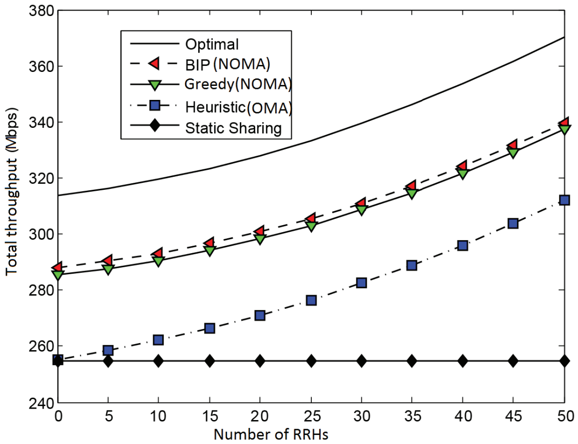

Figure 4. Total Throughput

VI. Conclusions

In this paper radio resource scheduling schemes for cloud radio access networks (C-RAN) utilising non-orthogonal multiple access (NOMA) is presented. Optimal and suboptimal solutions are provided. The solutions are compared with standard heuristic and static sharing schemes. The optimization solution is formulated as binary integer programming (BIP) problem which is NP-hard. The suboptimal scheme that gives solution to the optimization problem is derived. The scheme is derived by dividing the wireless resource allocation problem into sub-problems. Two schemes named BIP and greedy resource scheduling based on NOMA are derived. Performance evaluation of the proposed schemes are compared in terms of aggregate and total throughput. The simulation results show that the proposed schemes outperforms the standard heuristic and static sharing schemes in terms of throughput.

References

[1] H. Zhu, "Performance comparison between distributed antenna and microcellular systems," in *IEEE Journal on Selected Areas in Communications*, vol. 29, pp. 1151-1163, no. 6, 2011.

[2] J. Wang, H. Zhu, and N. Gomes, "Distributed antenna systems for mobile communications in high speed trains", in *IEEE Journal on Selected Areas in Communications*, vol. 30, pp. 675-683, May 2012.

[3] H Zhu, S Karachontzitis, and D Toumpakaris, "Low-complexity resource allocation and its application to distributed antenna systems," in *IEEE Wireless Communications*, vol. 17, pages 44-50, Issue 3, June 2010.

[4] H. Zhu and J. Wang, "Radio resource allocation in multiuser distributed antenna systems," in *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 10, pp. 2058-2066, October 2013.

[5] H. Zhu, "On frequency reuse in cooperative distributed antenna systems," in *IEEE Communications Magazine*, vol. 50, pages 85-89, Issue 4, April 2012.

[6] H. Osman, H. Zhu, D. Toumpakaris, and J. Wang, "Achievable rate evaluation of in-building distributed antenna systems," in *IEEE Transactions on Wireless Communications*, vol. 12, no. 7, pp. 3510-3521, July 2013.

[7] L. Liu, S. Bi, R. Zhang, "Joint power control and fronthaul rate allocation for throughput maximization in OFDMA-based cloud radio access network", in *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 4097-4110, Nov. 2015.

[8] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, Anxin Li, K. Higuchi, "Non-Orthogonal Multiple Access (NOMA) for Cellular Future Radio Access," in *IEEE 77th Vehicular Technology Conference (VTC Spring), 2013*, pp.1, 5, 2-5 June 2013.

[9] A. Benjebbour, Y. Saito, Y. Kishiyama, A. Li, A. Harada, and T. Nakamura, "Concept and practical considerations of non-orthogonal multiple access (NOMA) for future radio access," in *Proc. International Symposium on Intelligent Signal Processing and Communications Systems (ISPACS)*, 2013.

[10] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (noma)," in *Proc. IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, 2013.

[11] A. Benjebbour, A. Li, Y. Saito, Y. Kishiyama, A. Harada, and T. Nakamura, "System-level performance of downlink noma for future lte enhancements," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, 2013.

[12] J. Umehara, Y. Kishiyama, and K. Higuchi, "Enhancing user fairness in nonorthogonal access with successive interference cancellation for cellular downlink," in *Proc. IEEE International Conference on Communication Systems (ICCS)*, 2012.

[13] N. Otao, Y. Kishiyama, and K. Higuchi, "Performance of non-orthogonal access with sic in cellular downlink using proportional fair-based resource allocation," in *Proc. International Symposium on Wireless Communication Systems (ISWCS)*, 2012.

[14] H. Haci, H. Zhu, and J. Wang, "Performance of non-orthogonal multiple access (NOMA) with a novel asynchronous interference cancellation technique," in *IEEE Transactions on Communications*, vol. 65, pp. 1319-1335, issue 3, March 2017.

[15] H. Zhu and J. Wang, "Chunk-based resource allocation in OFDMA systems - Part I: chunk allocation," in *IEEE Transactions on Communications*, vol. 57, no. 9, pp. 2734-2744, Sept. 2009.

[16] H. Zhu and J. Wang, "Chunk-based resource allocation in OFDMA systems - Part II: joint chunk, power and bit allocation," in *IEEE Transactions on Communications*, vol. 60, no. 2, pp. 499-509, Feb. 2012.

[17] H. Zhu, "Radio resource allocation for OFDMA systems in high speed environments," in *IEEE Journal on Selected Areas in Communications*, vol. 30, pp. 748-759, May 2012.

[18] H. Zhu and J. Wang, "Performance analysis of chunk-based resource allocation in multi-cell OFDMA systems," in *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 2, pp. 367-375, February 2014.

[19] J. Choi, "Non-orthogonal multiple access in downlink coordinated two-point systems," in *IEEE Communications Letters*, 2014.

[20] R. Singh, "Sub-channel assignment and resource scheduling for non-orthogonal multiple access (NOMA) in downlink coordinated multi-point systems," in *2017 20th Conference on Innovations in Clouds, Internet and Networks (ICIN)*, Paris, 2017, pp. 17-22.

[21] M. Nair, Q. Z. Ahmed and H. Zhu, "Hybrid Digital-to-Analog Beamforming for Millimeter-Wave Systems with High User Density," in *2016 IEEE Global Communications Conference (GLOBECOM)*, Washington, DC, 2016, pp. 1-6.

[22] Vien, Q.-T., Ogbonna, N., Nguyen, H.X., et al., "Non-orthogonal multiple access for wireless downlink in cloud radio access networks", in *Proc. IEEE EW 2015*, Budapest, Hungary, May 2015, pp. 434–439.