IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 3, 2019, accepted December 4, 2019, date of publication December 17, 2019, date of current version January 3, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2960490

Energy Efficient SCMA Supported Downlink Cloud-RANs for 5G Networks

LILATUL FERDOUSE[1], SERHAT ERKUCUK[2], (Senior Member, IEEE), ALAGAN ANPALAGAN[1], (Senior Member, IEEE), AND ISAAC WOUNGANG[3]

[1]Department of Electrical and Computer Engineering, Ryerson University, Toronto, ON M5B 2K3, Canada
[2]Department of Electrical-Electronics Engineering, Kadir Has University, 34083 Istanbul, Turkey
[3]Department of Computer Science, Ryerson University, Toronto, ON M5B 2K3, Canada

Corresponding author: Alagan Anpalagan (alagan@ee.ryerson.ca)

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada, and in part by the Scientific and Technological Research Council of Turkey (TUBITAK).

ABSTRACT Cloud-radio access networks (C-RANs) are regarded as a promising solution to provide low cost services among users through the centralized coordination of baseband units for 5G wireless networks. The coordinated multi-point access, visualization and cloud computing technologies enable C-RANs to provide higher capacity and wider coverage, as well as manage the interference and mobility in a centralized coordinated way. However, C-RANs face many challenges due to massive connectivity and spectrum scarcity. If not properly handled, these challenges may degrade the overall performance. Recently, the non-orthogonal multiple access (NOMA) scheme has been suggested as an attractive solution to support multi-user resource sharing in order to improve the spectrum and energy efficiency in 5G wireless networks. In this paper, among various NOMA schemes, we consider and implement the sparse code multiple access (SCMA) scheme to jointly optimize the codebook (CB) and power allocation in the downlink of C-RANs, where the utilization of SCMA in C-RANs to improve the energy efficiency has not been investigated in detail in the literature. To solve this NP-hard joint optimization problem, we decompose the original problem into two sub-problems: codebook allocation and power allocation. Using the conflict graph, we propose the throughput aware SCMA CB selection (TASCBS) method, which generates a stable codebook allocation solution within a finite number of steps. For the power allocation solution, we propose the iterative level-based power allocation (ILPA) method, which incorporates different power allocation approaches (e.g., weighted and NOMA successive interference cancellation (SIC)) into different levels to satisfy the maximum power requirement. Simulation results show that the sum data rate and energy efficiency performances of SCMA supported C-RANs depend on the selected power allocation approach. In terms of energy efficiency, the performance significantly improves with the number of users when the NOMA-SIC aware geometric water-filling based power allocation method is used.

INDEX TERMS 5G wireless networks, C-RAN, NOMA, SCMA, power allocation, codebook assignment.

I. INTRODUCTION

The ever increasing number of mobile users, smart devices and applications increases the device density and mobile traffic loads, and eventually makes the traditional cellular networks incapable of handling such high demand. Therefore, researchers and engineers, both in academia and industry, have been working on new technological and architectural solutions for the next generation cellular networks, (e.g., in 5G) to handle 100 times more traffic and user loads, 1000 times higher network capacity and 1 ms latency. Moreover, spectrum efficiency (SE) and energy efficiency (EE) are the main focus of future wireless networks [1]. The cloud radio access network (C-RAN) is envisioned as a promising solution to utilize small cells in cellular networks along with macro cells to improve the total network capacity due to the spatial frequency reuse achieved by the flexible centralization of small cells through cloud computing [2].

The C-RAN consisting of remote radio heads (RRHs), the centralized baseband unit (BBU) pool and optical

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

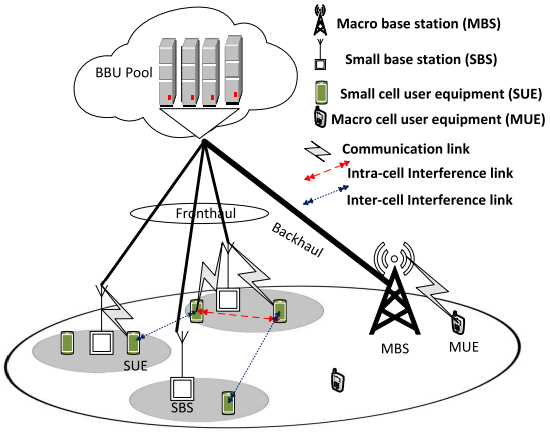


FIGURE 1. Two-tier heterogeneous cloud radio access network (H-CRAN).

fronthaul links is regarded as a new architectural paradigm that provides centralized and coordinated multi-point (CoMP) processing solutions to achieve higher capacity, SE, EE, seamless coverage, network control and cost efficient operation [3], [4]. In C-RANs, the macro base stations (MBSs) provide seamless coverage and network control, whereas the small cell base stations, regarded as RRHs, provide users with high data rates while leaving the basic control operations such as interference and handover management to macro cells. In addition, all the BBUs in a BBU pool are considered as computing servers to perform the baseband signal processing through cloud computing technologies. Moreover, the heterogeneous C-RAN architecture is considered as a good platform to manage the radio resource and interference in multi-cell environment. The BBU pool centrally coordinates the performance (e.g., EE and SE) of the entire network utilizing high-level cooperative interference cancellation techniques among small cells [5]. In Fig. 1, a two-tier heterogeneous cloud radio access network (H-CRAN) is depicted.

It is envisioned that the 5G mobile communication networks will support 100 times more connected devices per unit area compared to 4G LTE networks. Currently, the LTE and LTE-A networks support the orthogonal multiple access (OMA) technique, which utilizes a limited number of orthogonal resources for the users. Similarly, OMA supported H-CRANs utilize a limited number of orthogonal resources considered as communication resources for the small cell users [6]. To improve the spectrum efficiency in OMA supported H-CRANs, authors in [6] and [7] have considered an underlaid approach of orthogonal resources which are shared by both macro cell and small cell users. However, this underlaid approach is shown to increase the intra-cell and inter-cell interference levels, which limit the data rate of users. Considering the interference issues, authors in [6] have proposed an auction based distributed resource allocation with the aim to improve the data rate of small cell users. However, the OMA approach supports a limited number of connections due to the use of orthogonal resources. Therefore, to increase the connections per unit area, the non-orthogonal multiple access (NOMA) approach has been identified as a promising solution for future networks [8]. Unlike OMA, the NOMA methods utilize different power levels or user-specific signatures to provide the services to multiple users. In this context, the sparse code multiple access (SCMA), which is classified as one type of NOMA methods, assigns different codebooks (CBs) to different users. SCMA is regarded as a generalized low density signature (LDS) method, which uses sparse spreading sequences and SCMA communicates using sparse codewords in the codebooks. Furthermore, each codeword is composed of non-orthogonal resources such as sub-carriers (SCs) that are shared by different users [9]. The multiplexed signals of different users superimposed over the same sub-carrier can be decoded by the message passing algorithm [10] with low complexity. As the other NOMA type, power domain NOMA (PD-NOMA) method uses different power levels for multiple users to provide services in the same sub-carrier/subchannel and time slot [8]. Whether code domain or power domain, the NOMA methods support massive connectivity with better bandwidth utilization and provide higher SE and EE. However, the non-orthogonality in NOMA increases the mutual interference levels, therefore, a successive interference cancelation (SIC) method is applied at the receiver side [11].

NOMA in heterogeneous C-RAN architecture has been recently considered in [12]. Implementing NOMA in C-RAN may bring the advantages of SE, EE and massive connectivity through the centralized coordination in a BBU pool. However, the technical challenges of NOMA in the context of bandwidth and power allocation in C-RAN have not been investigated in detail in [12] nor in the literature. In this paper, we investigate in detail the SCMA method for C-RANs in terms of codebook and power allocation. SCMA has been chosen as the NOMA method as it can achieve a better sum data rate performance than PD-NOMA for the same signal-to-noise ratio, number of users and power conditions [13], [14]. While the SCMA method will be investigated in detail, its performance will be compared to both the PD-NOMA and OMA methods.

A. RELATED WORKS

The SCMA and PD-NOMA methods have been extensively studied in resource allocation problems in single cell networks [15]–[18] and multi-cell networks [20], [21]. A summary of these studies and the proposed solutions are given in Table 1. The resource allocation in PD-NOMA systems mainly considers power and subcarrier allocation to improve the system efficiency in terms of SE and EE. Similarly, in SCMA supported networks, the subcarrier allocation is referred to as CB or SC allocation for each user [22]. However, the joint optimization of power and SC allocation in PD-NOMA and SCMA systems is an NP-hard problem [18]. Therefore, the matching theory, greedy algorithm and auction methods are preferred for practical implementation. In [18], the authors consider an uplink single cell NOMA system and

TABLE 1. Summary of resource optimization problems.

Ref.	Network scenario	Link scenario	Multiple access		Scheduling			Power allocation	Solution approach
			SCMA	PD-NOMA	SC	CB	user		
[15]	-Single-cell	-Downlink	✓		✓			✓	Remove and reallocate iterative algorithm
[16]	-Single-cell	-Uplink		✓			✓		Optimal user grouping
[17]	-Single-cell	-Uplink	✓		✓				Matching theory
[18]	-Single-cell	-Uplink		✓	✓			✓	Matching theory, iterative water filling
[19]	-Single-cell	-Downlink		✓	✓			✓	Matching theory, DC programming
[20]	-Multi-cell	-Downlink		✓				✓	Distributed approach
[21]	-Multi-cell	-Downlink	✓	✓	✓	✓		✓	SCALE and GP
Proposed	-Multi-cellC-RAN	-Downlink	✓		✓	✓		✓	Conflict graph based TASCBS, ILPA

allocate the SCs to the users using geometric programming (GP) and a many-to-many matching model. For the power allocation, iterative water-filling (IWF) algorithm is applied to improve the data rate of the overall system. Similarly, a matching theory based power and SC allocation in downlink single cell networks is proposed in [19] and [23]. The energy efficient subchannel and power allocation for the NOMA method is considered in [19] and [24] where the authors in [19] have considered a single cell downlink network. A two-sided matching theory based subchannel allocation and difference of convex (DC) programming based power allocation method is proposed in [19]. Two-tier heterogeneous network (HetNet) with downlink power allocation is studied in [20], where the users receive the data from multiple access points using the CoMP NOMA method. Power budget and channel gain based SIC constraints are considered in the power optimization problem. For a downlink HetNet, the performances of PD-NOMA and SCMA methods are compared in [21] in terms of maximizing the data rate. Joint SC allocation and power optimization in PD-NOMA with the maximum power and SIC constraints are considered, whereas in the SCMA method, joint codebook and power optimization problem is considered only with the maximum base station power and SC sharing constraints. To solve the nonconvex joint optimization problem in PD-NOMA and SCMA, the authors utilize GP and successive convex approximation for low complexity (SCALE) algorithm which involves a series of convex relaxations [25]. Different from the above works, in this paper, we investigate the energy efficiency aspects of the SCMA method in C-RANs in terms of codebook and power allocation. In the resource allocation problem, the C-RAN small cell base station power budget, the fronthaul capacity constraint and the quality of service (QoS) requirement of each user, are considered.

B. CONTRIBUTIONS

In this work, we apply the conflict graph theory and geometric water filling approach to solve the codebook and power allocation in SCMA supported C-RANs with the objective to maximize the energy efficiency. The main contributions of this paper are as follows:

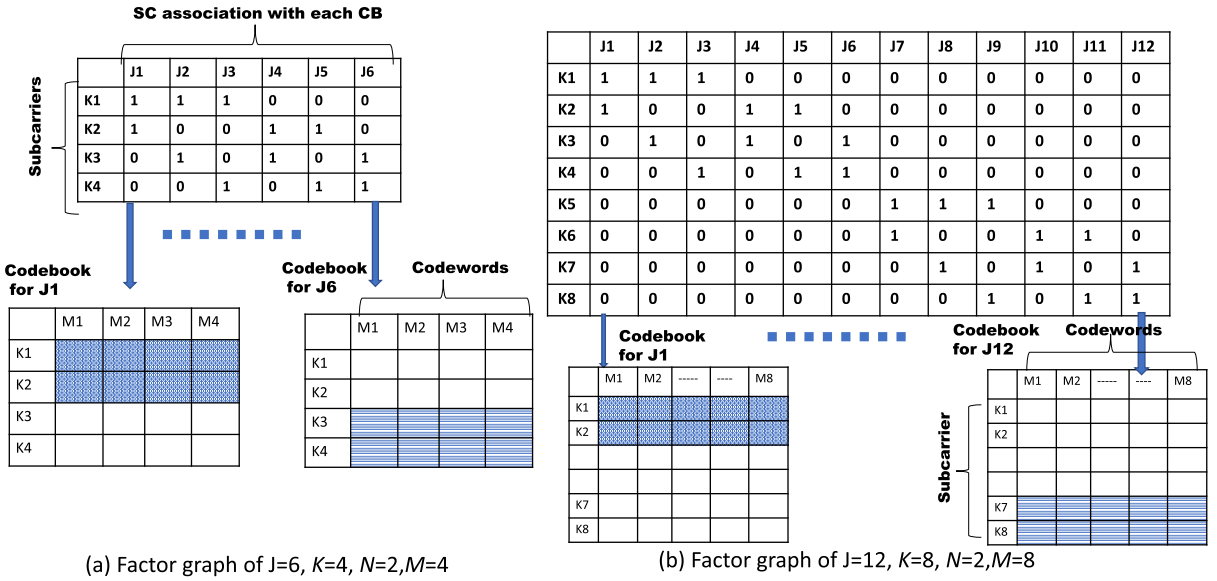
- The SCMA method is implemented to jointly optimize the codebook and power allocation in the downlink of the C-RANs. The structure of the SCMA codebook assignment and the optimization formulation of codebook and power allocation with the objective to improve the energy efficiency in C-RAN are presented in detail.
- To solve the joint optimization problem, the original problem is decomposed into two optimization problems. The first optimization problem is the codebook allocation (CA) with equal power allocation and the second optimization problem is the power allocation (PA) with the known codebook allocation.
- For the codebook allocation problem, the throughput aware SCMA codebook selection (TASCBS) method is proposed using the conflict graph theory. It is proven that the TASCBS method generates a stable codebook allocation solution within a finite number of steps.
- For the power allocation problem, the iterative level-based power allocation (ILPA) method, which incorporates different power allocation approaches (e.g., weighted and NOMA-SIC) into different levels to satisfy the maximum power requirement, is proposed.
- It is shown that the NOMA-SIC aware power allocation can be used with the geometric water filling method in the subcarrier level in a computationally efficient way and achieve higher energy efficiency compared to other power allocation approaches and the conventional PD-NOMA method.

The rest of the paper is organized as follows. In Section II, the SCMA codebook design and the system model for codebook and power allocation in C-RAN are described. The corresponding optimization problem and the proposed solutions for codebook allocation and power allocation are described in Sections III, IV and V, respectively. Section VI presents the numerical results of the proposed method. Finally, Section VII concludes the paper.

II. SYSTEM MODEL AND ASSUMPTIONS

A. SCMA CODEBOOK

We assume that the SCMA scheme supports a total number of K subcarriers, indexed by $k = \{1, 2, 3, \ldots, K\}$, where each



(a) Factor graph of J=6, *K*=4, *N*=2,*M*=4

(b) Factor graph of J=12, *K*=8, *N*=2,*M*=8

FIGURE 2. Factor graph and codebook structure.

TABLE 2. Relationship between number of SCs (*K*), number of codebooks (*J*), nonzero SCs (*N*) in *J* and the overloading factor (α).

	$K=4$, $N=2$	$K=8$, $N=2$	$K=12$, $N=2$	$K=16$, $N=2$
Without α $J=\frac{K!}{N!(K-N)!}$	$J=6$	$J=28$	$J=66$	$J=120$
With $\alpha=1.5$ $J=\alpha K$	$J=6$	$J=12$	$J=18$	$J=24$

codeword uses N out of K SCs. The total number of J codebooks are available in each cell, indexed by $j=\{1,2,...,J\}$, where $J \leq \frac{K!}{N!(K-N)!}$ and M is the total number of codewords in a codebook. The total number of codebooks is generally limited by the overloading factor $\alpha = \frac{J}{K}$. In the SCMA scheme, d_f denotes the number of users that can use the same subcarrier and d_v denotes the number of SCs used in each codeword. The relationships between SCs, codebooks and the overloading factor are shown in Table 2. The SC assignment in each SCMA codebook is represented by a factor graph $F=[f_{k,j}]$, where $f_{k,j}$ is the binary variable for SC allocation in each CB, represented as follows:

$$f_{k,j}=\begin{cases}1, & \text{if SC } k \text{ is assigned to codebook } j,\\ 0, & \text{otherwise.}\end{cases} \quad (1)$$

The CB design follows two constraints: 1) $\sum_{j=1}^{J} f_{k,j} \leq d_f, \forall k$ and 2) $\sum_{k=1}^{K} f_{k,j} \leq d_v, \forall j$. For simplicity, we assume that the base station knows the design of the codebook. Fig. 2 shows the factor graph representation and codebook assignments for two cases of J, K, M and N values. Using the factor graph, two sets of information, i.e., $\{S_{J_j}\}$ and $\{S_{K_k}\}$ can be retrieved. $\{S_{J_j}\}$ contains the set of non-zero SC information, which belongs to codebook j, i.e., $f_{k,j}=1$ and $k \in \{S_{J_j}\}$. Similarly, $j \in \{S_{K_k}\}$ contains the information about the codebooks, where the SC k is utilized. For example, In Fig. 2(a), $\{S_{J_1}\}=\{K_1,K_2\}, \{S_{J_2}\}=\{K_1,K_3\}, ..., \{S_{J_6}\}=\{K_3,K_4\}$ and $\{S_{K_1}\}=\{J_1,J_2,J_3\}, \{S_{K_2}\}=\{J_1,J_4,J_5\}, \ldots, \{S_{K_4}\}=\{J_3,J_5,J_6\}$. The base station assigns one codebook to one user with the goal to optimize the downlink codebook allocation and power allocation jointly to improve the system efficiency in terms of data rate and power. The list of symbols and parameters used in this paper is given in Table 3.

B. SYSTEM MODEL

We consider an SCMA based two-tier C-RAN, as shown in Fig. 1, where B is the total number of small cell base stations (SBSs), indexed by $b = \{1,2,....B\}$, which are covered by a single macro cell. Each SUE is equipped with one antenna and each small cell base station has ϕ antennas. The system supports a total number of small cell users $U = \sum_{b=1}^{B} U_b$, where a SUE in SBS b is indexed by $u=\{1,2,3,\ldots,U_b\}$. Let $Q=[q^b_{u,j}]$ be the CB association matrix, where $q^b_{u,j}$ is the binary variable for codebook allocation given by

$$q^b_{u,j}=\begin{cases}1, & \text{if codebook } j \text{ is assigned to SUE } u\\ & \text{on SBS } b,\\ 0, & \text{otherwise.}\end{cases} \quad (2)$$

The channel gain from SBS b to SUE u on SC k is denoted as $h^b_{u,k} \in \mathcal{C}^{\phi \times 1}$. The power allocated from SBS b to user u on SC k is denoted as $P^b_{u,k} \in (0, P^b_{max}]$, where P^b_{max} denotes the maximum power of SBS b. The signal-to-interference-plus noise ratio (SINR) achieved by SUE u, connected to SBS b

TABLE 3. List of symbols.

	Symbol	Description
Set	J	Total number of CBs
	K	Total number of SCs
	$\mathcal{B}$	Total number of SBS
	$\mathcal{U}$	Total number of SUEs
	$\mathcal{U}_b$	Total number of SUEs in SBS b
	M	Total number of codewords in each codebook
	S_{J_j}	Set of SCs in CB j
	S_{K_k}	Set of CBs using SC k
Index	k	Indexing for SC
	j	Indexing for CB
	b	Indexing for SBS
	u	Indexing for SUE
	U_u	User notation
	J_j	CB notation
	K_k	SC notation
Optimization Parameters	$f_{k,j}$	SC association parameter of k^{th} SC for j^{th} CB
	$q^b_{u,j}$	CB-User mapping parameter of j^{th} CB to u^{th} SUE on SBS b
	$P^b_{u,j}$	Power allocation from b^{th} SBS to u^{th} SUE on CB j
	$P^b_{u,k}$	Power allocation from b^{th} SBS to u^{th} SUE on SC k
Channel Parameters	$h^b_{u,k}$	The channel gain from b^{th} SBS to u^{th} SUE on k^{th} SC
	$h^b_{u,j}$	The channel gain from b^{th} SBS to u^{th} SUE on j^{th} CB
	P^b_{max}	Maximum power of SBS b
	$\gamma^b_{u,j}$	SINR of u^{th} SUE connected to b^{th} SBS on j^{th} CB
	$\Gamma_{k,u}$	SNR of u^{th} SUE on k^{th} SC
	$g_{k,u}$	Channel gain to noise ratio of u^{th} SUE on k^{th} SC
Others	N	Total number of nonzero elements in each SC
	F	Factor graph of CB design
	α	Overloading factor
	d_f	Total number of users for each SC
	d_v	Total number of SCs for each CB
	η	Energy efficiency
	$r^b_{u,j}$	Data rate of SUE u for b^{th} SBS on j^{th} CB
	$\tilde{r}_{u,j}$	Normalized rate of u^{th} SUE on j^{th} CB
	R_T	Total achievable data rate of all SUEs
	P_T	Total allocated power of all SBS

on codebook j can be written as

$$\gamma^b_{u,j} = \frac{q^b_{u,j}|h^b_{u,j}|^2 P^b_{u,j}}{I^b_{u,j}+\sigma^2}, = \frac{q^b_{u,j}\sum_{k\in\{S_{J_j}\}} f^b_{k,j}|h^b_{u,k}|^2 P^b_{u,k}}{I^b_{u,j}+\sigma^2}, \tag{3}$$

where σ^2 is the noise power. The channel gain and power allocation from SBS b to SUE u on CB j are denoted as $h^b_{u,j}$ and $P^b_{u,j}$, respectively. Here, we denote $P^b_{u,j}$ as CB-user power allocation, which is the summation of the SC-user power allocation obtaibed as

$$P^b_{u,j} = \sum_{k\in\{S_{J_j}\}} f^b_{k,j} P^b_{u,k}. \tag{4}$$

Similar to [13], [14], $I^b_{u,j}$ is the aggregated interference power, defined as

$$I^b_{u,j} = \underbrace{\sum_{u'\in U_b}\sum_{k\in\{S_{J_j}\}}\sum_{\substack{j'\neq j,\\ j'\in\{S_{K_k}||h^b_{u',k}|^2>|h^b_{u,k}|^2\}}} q^b_{u',j'} f^b_{k,j'}|h^b_{u,k}|^2 P^b_{u',k}}_{\text{Intra-cell interference}} + \underbrace{\sum_{b'=1,b'\neq b}^{B}\sum_{u'\in U_{b'}}\sum_{\forall j}\sum_{\forall k} q^{b'}_{u',j} f^{b'}_{k,j}|h^{b'}_{u',k}|^2 P^{b'}_{u',k}}_{\text{Inter-cell interference}}.$$

The first term on the right hand side is the intra-cell interference signal coming from other active SUEs (i.e., $u' \in U_b$) in the same cell, utilizing the same SCs (i.e., $k \in \{S_{J_j}\}$) in other codebooks, $j' \neq j$ where $j' \in \{S_{K_k}||h^b_{u',k}|^2 > |h^b_{u,k}|^2\}$. Similarly, the second term denotes the inter-cell interference coming from other cells utilizing the same SCs.

III. ENERGY EFFICIENT CODEBOOK ALLOCATION AND POWER ALLOCATION IN C-RAN

The objective of resource allocation in C-RANs is to maximize the EE in terms of codebook and power allocation. The EE can be measured by the total achievable data rate divided by the total allocated power (bits/J), written as

$$\eta = \frac{R_T}{P_T}. \tag{5}$$

According to the Shannon formula, the achievable data rate achieved by SUE u connected to SBS b using codebook j will be $r^b_{u,j} = \log_2(1+\gamma^b_{u,j})$, and the total data rate of all SUEs can be expressed as

$$R_T = \sum_{b\in B}\sum_{u\in U_b}\sum_{j\in J} r^b_{u,j}.$$

The total allocated power of small cells is denoted by

$$P_T = \underbrace{\sum_{b\in B}\sum_{u\in U_b}\sum_{j\in J} P^b_{u,j}}_{\text{dynamic}} + \underbrace{\sum_{b\in B} P^b_s}_{\text{static}}.$$

Similar to [26], we assume that each small cell has dynamic and static power factors. The dynamic power depends on the codebook allocation, whereas the circuit power is regarded as the static power P^b_s.

The mathematical formulation of the joint CA and PA in an SCMA based C-RAN system can be described as follows:

$$(P1)\ \max_{Q,P}\ \eta$$

subject to:

$$\text{C1:}\ \sum_{j\in J} q^b_{u,j} = 1, \quad \forall u\in U_b, b\in B,$$

$$\text{C2:}\ \sum_{u\in U_b} q^b_{u,j} = 1, \quad \forall j\in J, b\in B,$$

$$\text{C3:}\ q^b_{u,j}\sum_{j=1}^{J} P^b_{u,j} \le P^b_{max}, \quad \forall u\in U_b, b\in B,$$

$$\text{C4:}\ q^b_{u,j}\sum_{k\in\{S_{J_j}\}} f^b_{k,j}P^b_{u,k} \le P^b_{u,j}, \quad \forall u\in U_b, b\in B,$$

$$\text{C5:}\ r^b_{u,j} \ge r_{min}, \forall u\in U_b, b\in B,$$

$$\text{C6:}\ \sum_{u\in U_b} r^b_{u,j} \le R^b_{max}, \forall b\in B,$$

$$\text{C7:}\ q^b_{u,j}\in\{1,0\} \text{ and } P^b_{u,k}\ge 0. \tag{6}$$

In (6), the objective is to maximize the EE of the C-RAN system by allocating the same codebook among small cells

in an underlaid approach. Two optimization parameters are considered in this problem: i) the codebook allocation vector for small cell users (i.e., $q^b_{u,j} \in \{0,1\}$), and ii) the allocated power for small cell users (i.e., $P^b_{u,j}$ and $P^b_{u,k}$). C1 and C2 enforce that each SUE is connected to one SBS using one codebook. C3 ensures that the maximum power budget constraint for each small cell b, which is P^b_{max}, should be satisfied. C4 is the individual power budget constraint of each SUE and codebook in terms of SCs. C5 enforces the minimum data rate constraint of each user. C6 ensures the fronthaul capacity constraint in each SBS. R^b_{max} denotes the maximum fronthaul capacity of each SBS. C7 denotes the constraints on the binary codebook allocation and power allocation.

The objective function in (6) and the constraints C5 and C6 turn the problem **P1** into a mixed integer non-linear program (MINLP) with a non-convex feasibility set. The optimization problem **P1** is computationally intractable and NP-hard [27]. Hence, similarly as done in [28], [29], we have adopted a two-step iterative approach to solve the problem **P1**. We have split **P1** into two sub-problems: i) codebook allocation (**P2**) and ii) power allocation (**P3**). Assuming a fixed power allocation, the codebook allocation problem can be formulated as

$$\begin{aligned}(P2)\ \max_{Q}\ & R_T \\ \text{subject to: C1 to C2, C5, C6, and} & \\ \text{C7:}\quad & q^b_{u,j} \in \{1,0\}. \end{aligned} \tag{7}$$

Similarly, assuming a fixed codebook allocation the power optimization problem can be formulated as

$$\begin{aligned}(P3)\ \min_{P}\ & P_T \\ \text{subject to: C3 to C4, C5, C6,} & \\ \text{C7:} & P^b_{k,u} \geq 0. \end{aligned} \tag{8}$$

Initially, we address codebook allocation by assuming equal power allocation, no interference and minimum data rate constraints and propose the throughput aware SCMA CB Selection (TASCBS) method. Then, we adjust the power level using the channel gain to noise ratio, the effect of interference and minimum data rate information. The details of codebook and power allocation are explained in the following sections, respectively.

IV. CODEBOOK ALLOCATION

To solve the problem **P2**, we assume that the cloud controller of C-RAN knows all the channel state information (CSI) of the users and the factor graph of the CB design. For simplicity and without loss of generality, it is assumed that all the SBSs under the C-RAN use the same CB design. As for codebook allocation initialization, a bipartite graph is generated based on users' channel and rate information. The proposed CB allocation is explained as follows:

A. THROUGHPUT AWARE SCMA CB SELECTION (TASCBS)

Equal power allocation: According to the SCMA CB design, each user uses N number of non-zero SCs in each CB. Therefore, all the users receive an equal power from the BS according to the total number of CBs, i.e., $P_{u,j} = \frac{P^b_{max}}{|J|}$ and each SC receives an equal portion of the received power, i.e., $P_{u,k} = \frac{P_{u,j}}{|N|}$.

SNR estimation: Considering that there is fully available CSI information and no interference, the SNR of user u on SC k can be estimated as $\Gamma_{k,u} = g_{k,u}P_{u,k}$, where $g_{k,u} = \frac{|h_{u,k}|^2}{\sigma^2}$ represents the channel gain to noise ratio (CNR) defined as

$$CNR = \begin{array}{c} \\ K_1 \\ K_2 \\ . \\ . \\ K_K \end{array} \begin{array}{c} \begin{array}{ccccc} U_1 & U_2 & .. & .. & U_{U_b} \end{array} \\ \begin{pmatrix} g_{1,1} & g_{1,2} & .. & .. & g_{K_1,U_b} \\ g_{2,1} & g_{2,2} & .. & .. & g_{K_2,U_b} \\ .. & .. & .. & .. & .. \\ .. & .. & .. & .. & .. \\ g_{K_K,1} & g_{K_K,2} & .. & .. & g_{K_K,U_b} \end{pmatrix} \end{array}. \tag{9}$$

Normalized rate estimation: Based on the CNR and fixed CB design, the data rate of user u on codebook j can be estimated as $r_{u,j} = \log_2(1 + \sum_{k\in\{S_{J_j}\}} g_{k,u}P_{u,k})$. The normalized rates of each user can be obtained as $\tilde{r}_{u,j} = \frac{r_{u,j}}{\sum_{\forall j} r_{u,j}}$, where the summation of all normalized rates per user becomes one, i.e., $\sum_{\forall j} \tilde{r}_{u,j} = 1$. In (10), the normalized rate of each user for all codebooks are defined as

$$\begin{array}{c} \\ U_1 \\ U_2 \\ . \\ . \\ U_{U_b} \end{array} \begin{array}{c} \begin{array}{ccccc} J_1 & J_2 & .. & .. & J_J \end{array} \\ \begin{pmatrix} \{K_1,K_2\} & \{K_1,K_3\} & .. & .. & \{K_{..},K_{..}\} \\ \tilde{r}_{1,1} & \tilde{r}_{1,2} & .. & .. & \tilde{r}_{1,J} \\ \tilde{r}_{2,1} & \tilde{r}_{2,2} & .. & .. & \tilde{r}_{2,J} \\ .. & .. & .. & .. & .. \\ .. & .. & .. & .. & .. \\ \tilde{r}_{U_b,1} & \tilde{r}_{U_b,2} & .. & .. & \tilde{r}_{U_b,J} \end{pmatrix} \end{array}. \tag{10}$$

Bipartite graph and CB selection: A bipartite graph $G = \{U, J, E\}$ is depicted based on the users, CBs and the normalized rate information. The vertex set U denotes the set of users and the vertex set J represents the set of CBs. Each vertex of U is connected to vertex J based on the maximum normalized rate, i.e.,

$$E(u, j^*) = \max_{\forall j\in J} \tilde{r}_{u,j} \text{ and } \tilde{r}_{u,j} \geq r_{min}.$$

This means that each user u selects the CB j^* based on the maximum normalized rate. The flow graph of the throughput aware CB selection method is given in Fig. 3. The step-by-step procedure of the CB selection method is illustrated in Fig. 3. The CB selection and conflict resolution procedure are repeated until all the users are connected to their appropriate CBs. Fig. 4(a) shows an example of the bipartite graph representation for $U = 6$ and $J = 6$. When more than one user in the same SBS select the same CB $j^* \in J$, then j^* is represented as a conflict vertex in the bipartite

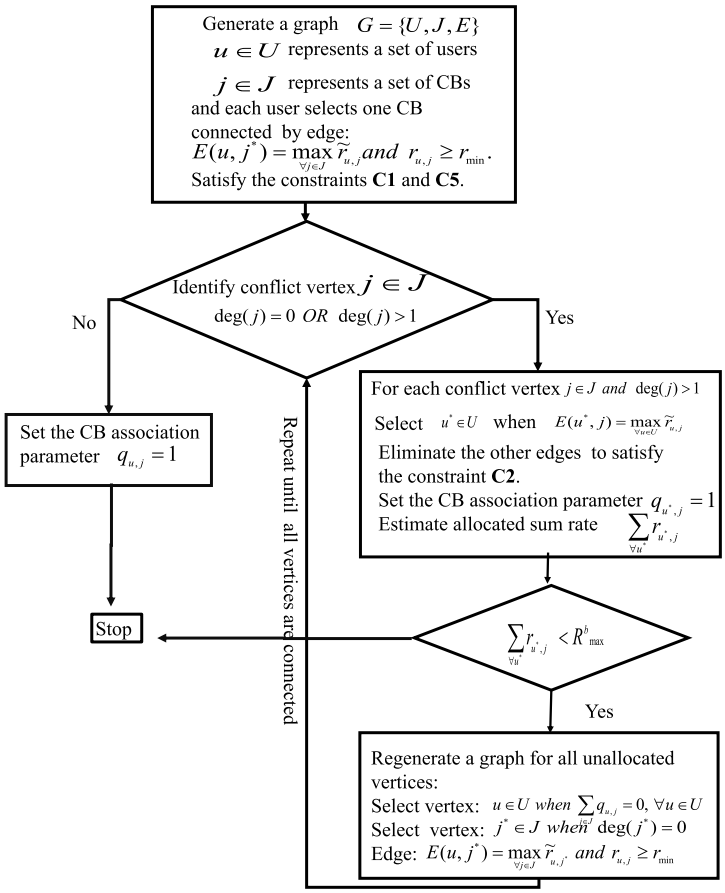


FIGURE 3. Throughput aware CB selection method.

graph. In Fig. 4(b), all the black (dark) color vertices in the CB represent the conflict vertex.

Definition 1: A vertex in $j \in J$ becomes a conflict vertex when it is matched by either more than one $u \in U$ or no u.

The conflict vertices ($j \in J$) in bipartite graph G are identified using the degree information of the vertices; i) $deg(j) = 0$ or, ii) $deg(j) > 1$. The conflict vertex violates the constraint **C2** in **P1** and **P2**. For conflict resolution, we propose a winner selection and edge elimination method, which is applied to all conflict vertices as long as each of the vertices in U is connected to exactly one vertex in J.

Conflict resolution: For the winner selection method, the vertex of $u^* \in U$ becomes the winner of j when its edge shows the maximum value among others. For example, conflict vertex $j \in J$, selects u^* when

$$E(u^*, j) = \max_{\forall u \in U} \tilde{r}_{u,j} \tag{11}$$

and eliminates other edges to satisfy the condition **C2** in **P2**. This process is repeated for all other conflict vertices which have $deg(j) > 1$. Fig. 4(c) shows the graphical representation after the winner selection and edge elimination method has been applied.

B. STABILITY AND CONVERGENCE OF THE TASCBS METHOD

Definition 2: A stable allocation is defined as no conflict vertex and each vertex in $u \in U$ is connected to at most one vertex in $j \in J$ and vice versa.

Lemma 1: The TASCBS converges to a pair-stable allocation, when the number of conflict vertices becomes zero.

Proof: According to the codebook selection method given in Fig. 3, when the proposed TASCBS converges to a stable allocation, no user $u \in U$ has a conflict with another user for the same choice of CB $j^* \in J$. If j^* is selected by more than one users, then the TASCBS resolves the conflict by picking up the best user, who will benefit the most by using the utility (11). Thus, the matches of user u^* must be the best choice for other users in the current situation. Hence, the terminal matching is pair-stable.

Theorem 1: The proposed TASCBS converges to a stable allocation after a limited number of iterations.

Proof: See **Appendix A.**

V. POWER ALLOCATION

In each cell[1], using the known values of the CB allocation parameter Q and the factor graph F, the power allocation

[1]Since the power allocation is performed separately at each SBS, without loss of generality, we drop the superscript b.

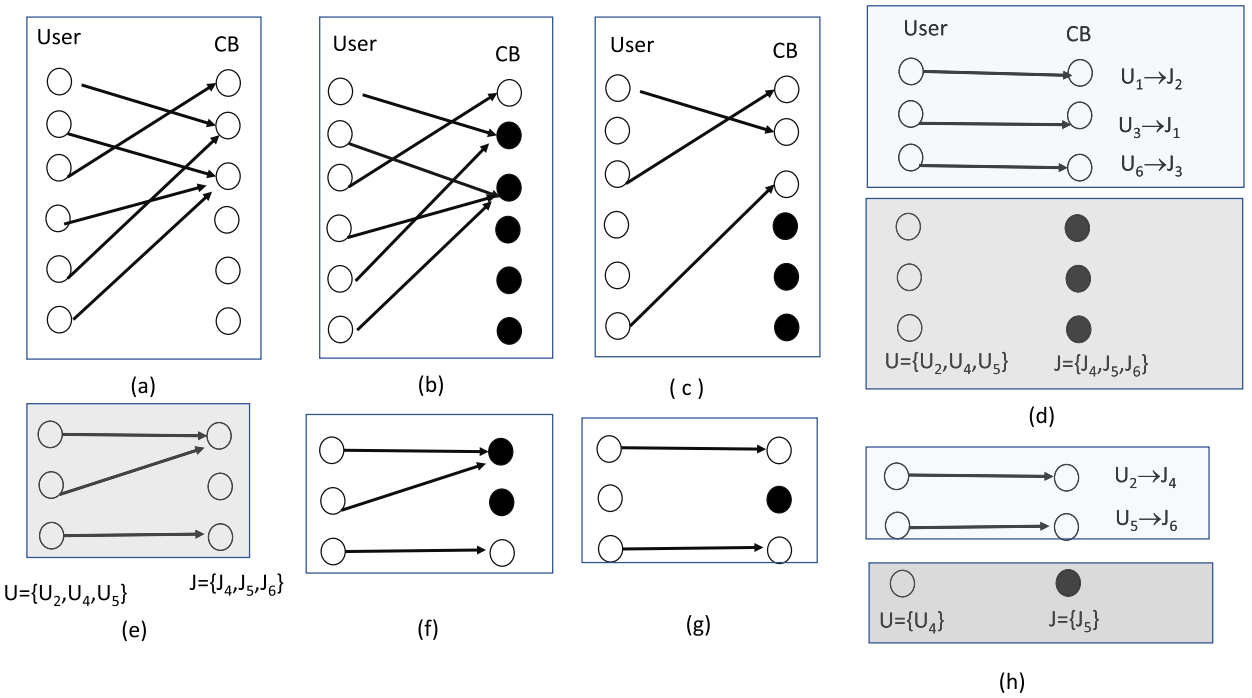


FIGURE 4. (a) Bipartite graph, (b) Identifying conflict vertices, (c) Winner selection and edge elimination, (d) Allocated and unallocated vertices, (e) Repeated case of (a), (f) Repeated case of (b), (g) Repeated case of (c), (h) Repeated case of (d).

problem (**P3**) can be reformulated as

$$
\begin{aligned}
(P3)\ \min_{P_{u,k}}\ & P_T \\
\text{subject to:}\ & \\
\text{C3:}\quad & P_{max} - \sum_{j=1}^{J} P_{u,j} \geq 0, \\
\text{C4:}\quad & P_{u,j} - \sum_{k \in \{S_{J_j}\}} P_{u,k} \geq 0, \\
\text{C5:}\quad & \log_2(1+\gamma_{u,j}) - r_{min} \geq 0, \\
\text{C6:}\quad & R_{max} - \sum_{u \in U_b} \log_2(1+\gamma_{u,j}) \geq 0, \\
\text{C7:}\quad & P_{u,k} \geq 0. \qquad (12)
\end{aligned}
$$

To perform the power allocation, we use Karush-Kuhn-Tucker (KKT) optimality and define the following Lagrangian function

$$
\begin{aligned}
&\mathbb{L}(P_{u,j}, P_{u,k}, \lambda, \beta, \phi, \psi) \\
&= \sum_{\forall u}\sum_{\forall j} P_{u,j} + P_s - \lambda\{P_{max} \\
&\quad - \sum_{\forall j} P_{u,j}\} - \sum_{\forall j}\beta_j \sum_{\forall u}\{P_{u,j} - \sum_{k \in \{S_{J_j}\}} P_{u,k}\} \\
&\quad - \sum_{\forall j}\phi_j \sum_{\forall u}\{\log_2(1+\gamma_{u,j}) - r_{min}\} \\
&\quad - \sum_{\forall j}\psi_j\{R_{max} - \sum_{\forall u}\log_2(1+\gamma_{u,j})\}, \qquad (13)
\end{aligned}
$$

where λ, β, ϕ are the Lagrange multipliers for the constraints C3-C5 of problem **P3**, respectively. Differentiating (13) with respect to $P_{u,k}$, we obtain the following power allocation of SUE u over SC k as

$$
P_{u,k} = \left[\frac{\phi_j - \psi_j}{\ln(1+\lambda+\beta_j)} - \frac{1}{\delta_{u,k}}\right]^+, \qquad (14)
$$

where $\delta_{u,k} = \frac{|h_{u,k}|^2}{I^b_{u,j}+\sigma^2}$ and $[\varepsilon]^+ = \max(\varepsilon, 0)$, which is a multi-level water filling allocation [27]. The calculation of (14) is provided in **Appendix B.**

In the SCMA scheme, each SC is shared among d_f number of users, therefore it is important to consider the intra-cell interference during the SC-user power allocation. On the other hand, the SC-user PA in (14) depends on the optimal choice of λ, ϕ, and δ values and do not consider the interference cancellation during PA. Thus, we propose an iterative level based PA, which exploits the relation between users, CBs, SCs and users' channel state information during the power allocation.

A. ITERATIVE LEVEL-BASED POWER ALLOCATION (ILPA)

The proposed iterative level-based power allocation (ILPA) method consists of five levels to solve the problem **P3**. The levels one (L1) to three (L3) are executed with the assumption of no interference and equal power allocation based on the CNR information. In levels four (L4) and five (L5), the power level of each SC is adjusted using the NOMA SIC principle [8]. The NOMA SIC method helps to mitigate the intra-cell interference, has been applied for SCMA [13], [14]. For inter-cell interference, we assume that the centralized cloud controller in the C-RAN applies the enhanced inter-cell interference cancellation (eICIC) method.

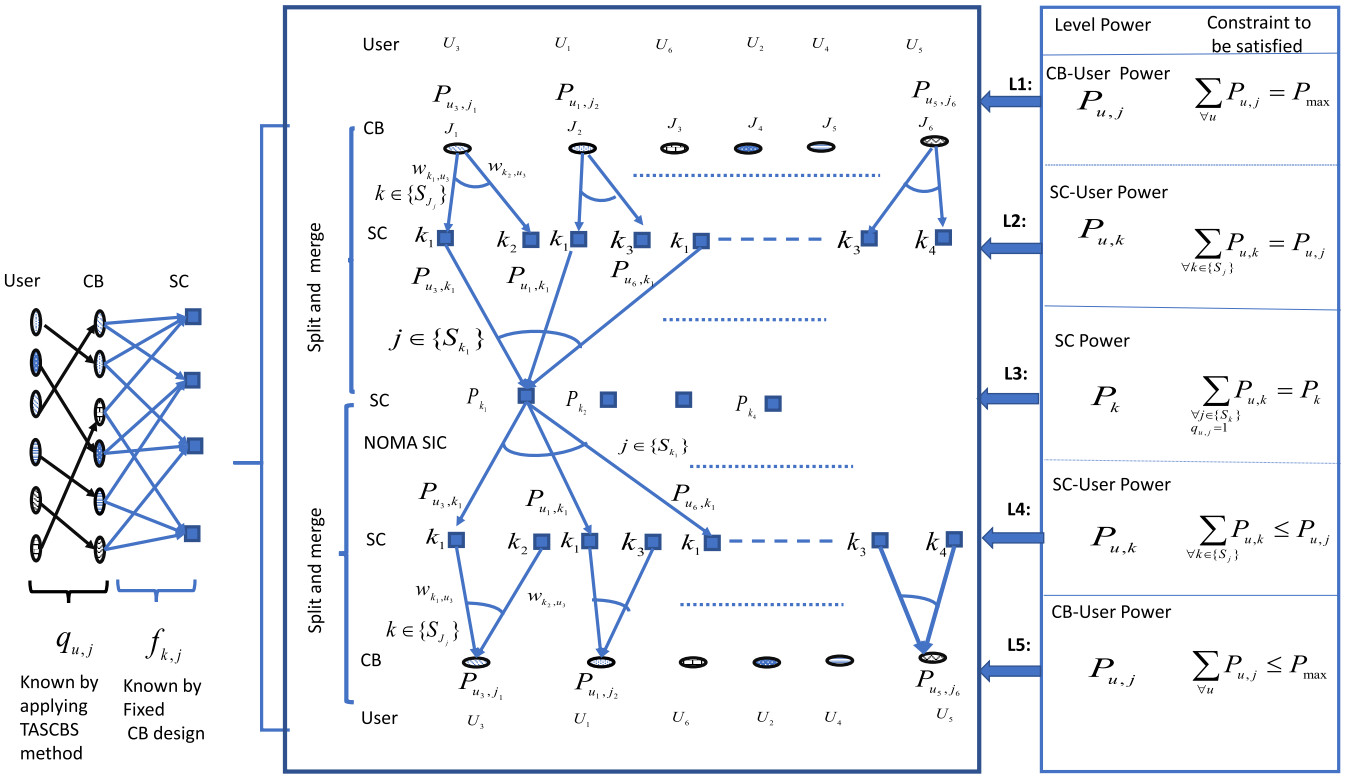


FIGURE 5. Iterative level-based power allocation in a single cell where $K = 4$, $N = 2$ and $J = 6$.

This eICIC method is used to mitigate the interference among the cells in heterogeneous networks, and it is easily adaptable in C-RAN environment with the advancement of software defined cloud controller [5].

The ILPA method works as follows:

L1: After the CB association parameter ($Q = [q^b_{u,j}]$) is obtained from CA, the CB-user power is equally divided based on the total number of CBs available in each SBS, i.e., $P_{u,j} = \frac{P_{max}}{|J|}$.

L2: The SC power for each user is allocated based on the CB design information ($F = [f_{k,j}]$), CB association parameter and CNR information. Suppose that the CB association parameters are obtained after executing the TASCBA method. An example is shown in Fig. 5 illustrating the case where the total number of SCs is $K = 4$, the nonzero elements of each SC is $N = 2$ and the total number of CBs is $J = 6$. Accordingly, the maximum number of users supported by the system is $U_b = 6$. Let the CB-user association matrix be as follows:

$$Q = \begin{array}{c} \\ U_1 \\ U_2 \\ U_3 \\ U_4 \\ U_5 \\ U_6 \end{array} \begin{array}{c} \begin{array}{cccccc} J_1 & J_2 & J_3 & J_4 & J_5 & J_6 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \end{array}.$$

According to the CB-user association matrix and the CNR information in (9), the weight of CNR for each SC in a CB is estimated as

$$w^j_{k,u} = \frac{g_{k,u}}{\sum_{\forall k \in \{S_{J_j}\}} g_{k,j}} \tag{15}$$

and it is associated with each SC as

$$\begin{array}{c} \\ U_1 \\ U_2 \\ U_3 \\ .. \\ U_6 \end{array} \begin{array}{c} \begin{array}{cccccc} J_1 & & J_2 & & .. \; .. & J_6 \end{array} \\ \begin{pmatrix} \{K_1, & K_2\} & \{K_1, & K_3\} & .. \; .. & \{K_3, & K_4\} \\ 0 & 0 & w^2_{1,1} & w^2_{3,1} & .. \; .. & 0 & 0 \\ 0 & 0 & 0 & 0 & .. \; .. & 0 & 0 \\ w^1_{1,3} & w^1_{2,3} & 0 & 0 & .. \; .. & 0 & 0 \\ .. & .. & .. & .. & .. \; .. & .. & .. \\ 0 & 0 & 0 & 0 & .. \; .. & .. & .. \end{pmatrix} \end{array}.$$

Similar to [21], we utilize the weight-based power allocation, where the summation of all SC weights of each CB is equal to one, i.e.,

$$\sum_{\forall k \in \{S_{J_j}\}} w^j_{k,u} = 1.$$

The power in **L2** is estimated as

$$P_{u,k} = P_{u,j} \times w^j_{k,u}. \tag{16}$$

L3: Each SC power is estimated as

$$P_k = q_{u,j} \sum_{j \in \{S_{\mathcal{K}_k}\}} P_{u,k}, \tag{17}$$

where

$$(P3.1)\ \min_{P_{u,k}} P_{u,j}$$
subject to:
$$\text{C4:} q_{u,j} \sum_{k\in\{S_{J_j}\}} P_{u,k} \le P_{u,j},$$
$$\text{C7:} P_{u,k} \ge 0. \quad (18)$$

$$(P3.2)\ \min_{P_{u,j}} P_T$$
subject to:
$$\text{C3:} \sum_{j=1}^{J} P_{u,j} \le P_{max},$$
$$\text{C7:} P_{u,j} \ge 0. \quad (19)$$

L4: In a single cell scenario, the problem **P3** is divided into two subproblems **P3.1** and **P3.2** as shown above. The problem **P3.1** is regarded as a SC-user power allocation, whereas **P3.2** is regarded as a CB-user power allocation. An iterative geometric water filling (GWF) method is applied to solve the problem **P3.1**. However, the main challenge is that when the SC power (P_k) is allocated among the users, they may not be causing a minimum interference to each other. Thus, we consider the NOMA SIC principle, where the SC power is allocated in such a way that the users with the highest CNR get the lowest power.

To implement this approach, we define the step depth in GWF in such a way that the highest CNR is represented with the highest depth, so that the lowest power is allocated to it. Therefore, we represent the step depth, $d_{k,u}$, as the inverse of the normalized CNR as

$$d_{k,u} = \frac{\sum_{j\in\{S_{K_k}\}} g_{k,u}}{g_{k,u}}. \quad (20)$$

Fig. 6 shows the NOMA SIC aware GWF procedure. Using (20), the inverse of the normalized CNR is represented as the largest depth, hence, we can apply the GWF method to find the explicit solution of (**P3.1**) in a computationally efficient way. According to [30], the explicit solution of (**P3.1**) is:

$$P_{u,k} = \begin{cases} P_{l^*,k} + (d_{l^*} - d_u), & \text{if } 1 < u \le l^* \\ 0, & l^* < u < |\{S_{\mathcal{K}_k}\}|. \end{cases} \quad (21)$$

Here, l^* denotes the maximum water level and P_{k,l^*} denotes the allocated power in l^* level.

L5: Finally, the user adjusts the power level to its associated CB as follows:

$$P_{u,j} = \sum_{k\in\{S_{J_j}\}} w^j_{k,u} \sum_{k\in\{S_{J_j}\}} P_{u,k}.c$$

A numerical example of the ILPA method is provided in **Appendix C**.

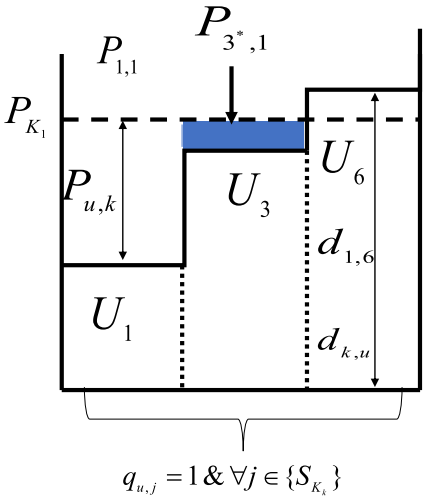


FIGURE 6. NOMA SIC aware GWF procedure.

TABLE 4. Simulation parameters.

Parameters	Values
Total no. of small cells, B	3
Total no. SUEs in each SBS, U_b	6
Total no. of SCs, K	4
Total no. of CBs, J	6
The nonzero elements in each codeword, N	2
Max no. of users in each SC, d_f	3
Max no. of SCs are used by each users, d_v	2
Factor graph F	$[1,1,1,0,0,0;$ $1,0,0,1,1,0;$ $0,1,0,1,0,1;$ $0,0,1,0,1,1]$
Radius of small cell	10 m
Minimum data rate requirements, r_{min}	50-140 kbps
Maximum power of SBS, P_{max}	30 dBm
Path-loss exponent	4
Noise power spectrum density	-144 dBm/Hz

VI. SIMULATION RESULTS

In this section, the sum data rate and EE performances of the proposed TASCBS and ILPA methods for SCMA supported downlink C-RANs are investigated. In the simulation model, we consider a 120m $\times$ 100m area, where one macro base station is underlaid by 3 small cell base stations. The locations of SBSs and SUEs are modeled using spatial Poisson point process (PPP) with predefined intensity values. For the CB assignment, we consider the $J = 6$, $K = 4$ and $N = 2$ case for each SBS. The simulation parameters are shown in Table 4. For performance evaluations, we compare SCMA with the PD-NOMA and OMA methods. In PD-NOMA, the users are grouped according to their locations and at most two users are share the same SC, which is helpful to reduce the error propagation of SIC [31], [20]. We initially consider an equal PA in each SC, and we compare the SCMA,

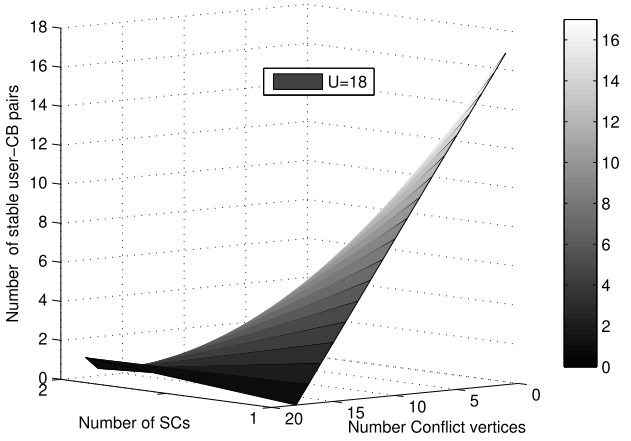


FIGURE 7. Number of stable user-CB pairs versus number of conflict vertices in TASCBS method, where $B = 3$, $J = 6$, $K = 4$ and $N = 2$, showing the convergence of TASCBS method.

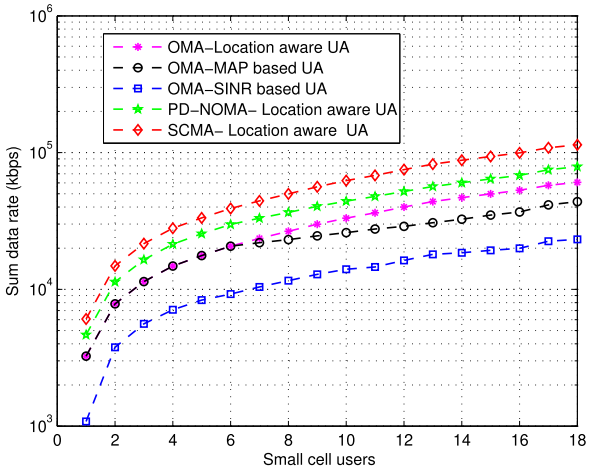


FIGURE 8. Performance of sum data rate with equal PA, where total number of SBS is $B = 3$ and each SBS supports $U_b = 6$ SUEs, $J = 6$ CBs, and $K = 4$ SCs.

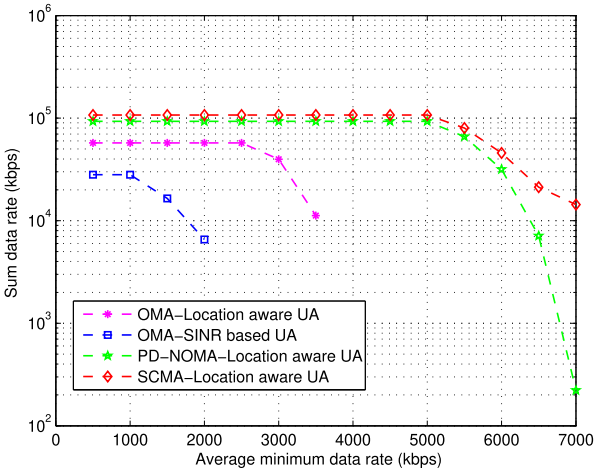


FIGURE 9. Performance of sum rate with different minimum data rate requirements.

PD-NOMA and OMA bandwidth allocation with different user association (UA) schemes. Then, we apply the ILPA method for the SCMA system and discuss the effect of PA on the EE of the C-RANs. For the bandwidth allocation in OMA, we consider an orthogonal SC allocation, whereas in the SCMA method, we consider a CB allocation using the TASCBS method. For the power allocation in PD-NOMA, we consider the mesh adaptive direct search (MADS) algorithm proposed in [32]. Three different UA schemes, namely, i) location aware, ii) SINR based, and iii) maximum a posteriori (MAP) based schemes [33], [34] are considered for comparison purpose.

Fig. 7 shows the convergence behavior of the TASCBS method. It is observed that when the number of conflict vertices becomes zero, the TASCBS method results in a stable CB assignment for all users. Here, we have considered $B = 3$ SBSs with $J = 6$ CBs in each SBS, where the total number of users in each SBS is $U_b = 6$ and the total number of users in the network is $U = 18$. Note that the TASCBS method is applied to each SBS separately. In Fig. 7, it can be observed that initially the number of conflict vertices becomes high. Gradually, the number of conflict vertices decreases when the number of user-CB pairs is increased. When the total number of stable user-CB pairs is 18, there are no more conflict vertices.

Fig. 8 shows the performance of the sum data rate versus the number of small cell users for different user association methods in OMA, PD-NOMA and SCMA, considering an equal PA in each SC. It can be observed that the sum data rate performance of SCMA using the location aware UA method gives the best result among all others. This can be explained by the nonorthogonal allocation of the CBs and the fact that the users do not cause significant interference in the same SC. As for SCMA and PD-NOMA comparison, SCMA can provide higher data rate as also reported in [13], [14] due to better accommodation of users. In OMA with the location-aware UA method, the users are associated with the closer proximity to the base station, showing a similar performance with respect to the MAP based method. Since the MAP based UA works on the maximum received CSI information, the users in close proximity to the base station receive the maximum CSI information.

A. PERFORMANCE OF SCMA WITH MINIMUM DATA RATE AND FRONTHAUL CAPACITY REQUIREMENTS

Fig. 9 shows the sum data rate performances among three methods with different minimum data rate requirements. Note that the minimum data rate requirement is an important factor for choosing the multiple access and user association methods. When the minimum data rate requirement is increased, the OMA-SINR based UA method can only support a fewer number of users compared to the location-aware UA method. For the SCMA location-aware UA, each user uses $d_v = 2$ SCs and each SC is shared by $d_f = 3$ users in a nonorthogonal way. Hence, it can accommodate an increased number of users while satisfying the minimum data rate constraint. Similarly, the PD-NOMA method also supports more users than the OMA method and the users

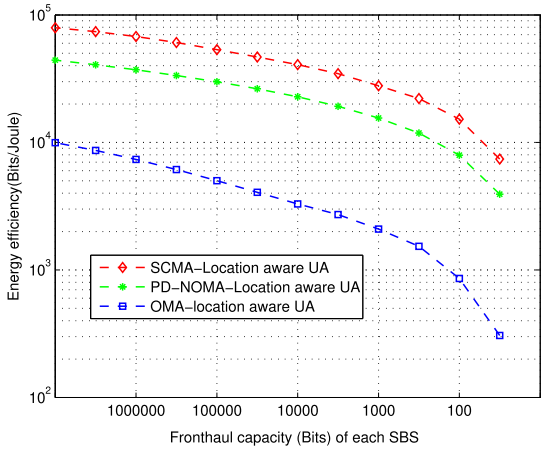


FIGURE 10. EE performance with different fronthaul capacity values.

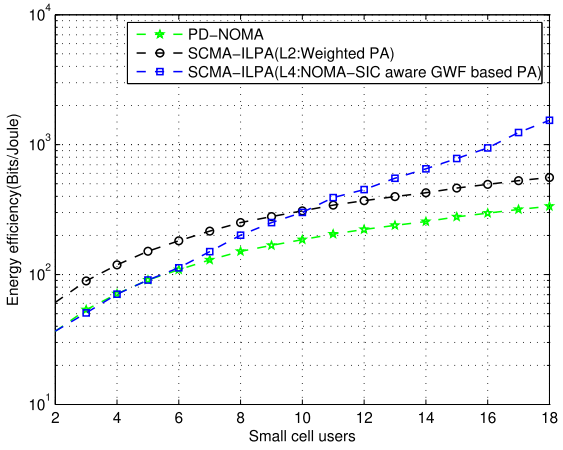


FIGURE 11. EE performance of SCMA with ILPA.

can share the SCs in a nonorthogonal way. On the other hand, in the OMA method with location-aware and the SINR-based UA methods, each user can choose only one orthogonal SC based on the relative distance or average received signal strength. Therefore, fewer users are supported and the sum data rate is lower compared to that obtained using SCMA.

Fig. 10 depicts the energy efficiency versus the fronthaul capacity of each SBS when the minimum data rate requirement is $r_{min} = 50$ kbps. It is shown that with decreasing fronthaul capacity of SBSs, the energy efficiency decreases. This is a result of less users being associated with a SBS to satisfy the minimum rate requirements when the capacity of each SBS is decreased. Consequently, there is less resource consumption in the C-RAN, which leads to lower EE with a decreasing fronthaul capacity in all multiple access methods.

B. PERFORMANCE OF SCMA WITH ILPA METHOD

The EE performance of the SCMA method is compared for two power allocation approaches in the ILPA method, namely: i) the weighted PA in L2, and ii) the NOMA-SIC aware GWF-based PA in L4. In the PD-NOMA method, the power constraint is satisfied by utilizing the MADS algorithm. The EE performance of L4 power allocation

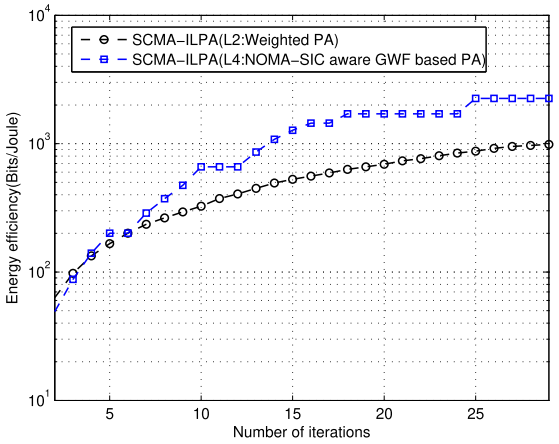


FIGURE 12. Convergence behavior of ILPA.

increases with the number of users and L4 PA shows better performance compared to the weighted PA and PD-NOMA as shown in Fig. 11. According to the CB design, each user uses 2 SCs and the weighted PA utilizes more power on the better channel. On the other hand, the level four in the ILPA method allocates the power to the SCs according to the NOMA SIC principle. According to this principle, among the 3 users in the same SC, the one which has the highest channel gain utilizes less power to avoid intra-SC interference, hence, the EE performance of this level becomes better than the weighted PA. Therefore, it can be concluded that the EE performance of SCMA becomes more significant when the NOMA SIC is incorporated into the SCMA PA for higher number of users. Finally, Fig. 12 shows the convergence behavior of the levels L2 and L4 of the ILPA method, where $J = 6$, $K = 4$ and $N = 2$. It is apparent from the figure that the weighted PA in L2 and NOMA-SIC aware GWF based PA in L4 schemes show non-decreasing energy efficiency and converge within a limited number of iterations.

VII. CONCLUSION

In this paper, we considered the EE performance of SCMA supported C-RANs. We implemented the SCMA method to jointly optimize the codebook and power allocation in the downlink of C-RANs. To solve the optimization problem, we proposed the throughput aware SCMA codebook selection and iterative level-based power allocation methods. From the implementation perspective, the software defined cloud controller executes the TASCBS at each SBS, which results in a stable codebook allocation solution within a finite number of steps. After obtaining the codebook allocation solution, the ILPA method helps to optimize the power allocation for the whole C-RAN. Simulation results show that the EE performance of NOMA SIC aware GWF based power allocation is better than that of the weighted power allocation scheme with increasing number of users. The sum data rate and EE performances mainly depend on dynamic codebook design, which is a future research topic of investigation. The proposed approaches TASCBS and ILPA are important for practical implementation of SCMA supported downlink C-RANs.

APPENDIX

A. PROOF OF THEOREM 1

In TASCBS, each user $u \in U$ selects a CB $j \in J$ without knowing other users' choices. This increases the possibility of a conflict, also named as a conflict vertex. Each iteration in TASCBS resolves the conflict vertices and eliminates the edges which violate the one-to-one matching criterion, i.e., **C1** and **C2** in **P2**. There are J CBs in each small cell, so the number of selections that each user u makes for the CB is no larger than J, and thus, the total number of iterations is no more than J. Also, after each iteration, users and CBs are categorized into two groups, i.e., allocated and unallocated. The TASCBS procedure is repeated for the unallocated groups until there are no conflict choices. Therefore, the TASCBS converges to a stable allocation according to Theorem 1.

B. CALCULATION OF POWER ALLOCATION

Using the relation of CB-user power and SC-user power in (4) and letting $\delta_{u,k} = \frac{|h_{u,k}|^2}{I^b_{u,j}+\sigma^2}$, the problem of (22) becomes

$$\begin{aligned}
&\mathbb{L}(P_{u,j}, P_{u,k}, \lambda, \beta, \phi, \psi) \\
&= \sum_{\forall u}\sum_{\forall j}\sum_{k\in\{S_{J_j}} P_{u,k} + P_s \\
&\quad - \lambda\left\{P_{max} - \sum_{\forall j}\sum_{k\in\{S_{J_j}\}} P_{u,k}\right\} \\
&\quad - \sum_{\forall j}\beta_j\sum_{\forall u}\left\{P_{u,j} - \sum_{k\in\{S_{J_j}\}} P_{u,k}\right\} \\
&\quad - \sum_{\forall j}\phi_j\sum_{\forall u}\left\{\log_2\left(1+\sum_{k\in\{S_{J_j}\}} P_{u,k}\delta_{u,k}\right) - r_{min}\right\} \\
&\quad - \sum_{\forall j}\psi_j\left\{R_{max} - \sum_{\forall u}\log_2\left(1+\sum_{k\in\{S_{J_j}\}} P_{u,k}\delta_{u,k}\right)\right\} \qquad (22)
\end{aligned}$$

Minimizing **P3** for any given Q and F is equivalent to differentiating $\mathbb{L}$ with respect to $P_{u,k}$ and setting the result to zero. That is

$$\frac{\partial \mathbb{L}}{\partial P_{u,k}} = 0$$

$$1 + \lambda + \beta_j - \frac{\phi_j\delta_{u,k}}{\ln(1+P_{u,k}\delta_{u,k})} + \frac{\psi_j\delta_{u,k}}{\ln(1+P_{u,k}\delta_{u,k})} = 0$$

Therefore, $P_{u,k}$ can be obtained as

$$P_{u,k} = \left[\frac{\phi_j - \psi_j}{\ln(1+\lambda+\beta_j)} - \frac{1}{\delta_{u,k}}\right]^+.$$

C. NUMERICAL EXAMPLE OF THE ILPA METHOD

To compute the power levels **L1**-**L5** in the ILPA method, we consider the case of Fig. 5, where the total number of SCs is $K = 4$, the nonzero elements of each SC is $N = 2$ and the total number of CBs is $J = 6$. Accordingly, the maximum number of users supported by the system is $U_b = 6$. Assume that the total maximum power budget of the base station is $P_{max} = 12$dB. Let the CB-user association matrix be

$$Q = \begin{array}{c} \\ U_1 \\ U_2 \\ U_3 \\ U_4 \\ U_5 \\ U_6 \end{array}\begin{array}{c} \begin{array}{cccccc} J_1 & J_2 & J_3 & J_4 & J_5 & J_6 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \end{array} \qquad (23)$$

and the CNR matrix be

$$CNR = \begin{array}{c} \\ K_1 \\ K_2 \\ K_3 \\ K_4 \end{array}\begin{array}{c} \begin{array}{cccccc} U_1 & U_2 & U_3 & U_4 & U_5 & U_6 \end{array} \\ \begin{pmatrix} .7 & .2 & .6 & .2 & .1 & .8 \\ .2 & .8 & .7 & .8 & .2 & .2 \\ .5 & .5 & .5 & .2 & .5 & .2 \\ .4 & .3 & .2 & .5 & .6 & .7 \end{pmatrix} \end{array}. \qquad (24)$$

According to the CB-user association matrix in (23) and the CNR matrix in (24), equation (15) can be applied to estimate the weight of CNR for each SC in a CB. The computation results are given in (25), as shown at the bottom of this page.

L1: The power levels in L1 are estimated as $P_{u,j} = \frac{P_{max}}{|J|} = \frac{12}{6} = 2$dB. According to the CB-user association matrix in (23), the CB-user power is $P_{U_1,J_2} = P_{U_2,J_4} = P_{U_3,J_1} = P_{U_4,J_5} = P_{U_5,J_6} = P_{U_6,J_3} = 2$dB.

L2: Using the weighted SC-user power in (16), the power levels in L2 can be computed, where they satisfy the constraints C3 and C4 in **P1** as shown in Table 5.

L3: Using (17), the SC powers can be computed as follows:

$$P_{K_1} = P_{3,1} + P_{1,1} + P_{6,1} = .92 + 1.16 + 1.06 = 3.14\text{dB}$$

$$P_{K_2} = P_{3,2} + P_{2,2} + P_{4,2} = 1.08 + 1.24 + 1.24 = 3.56\text{ dB}$$

$$P_{K_3} = P_{1,3} + P_{2,3} + P_{5,3} = .84 + .76 + .9 = 2.5\text{dB}$$

$$P_{K_4} = P_{6,4} + P_{4,4} + P_{5,4} = 0.94 + 0.76 + 1.1 = 2.8\text{dB}$$

$$P_{K_1} + P_{K_2} + P_{K_3} + P_{K_4} = 12\text{dB}$$

$$\begin{array}{c} \\ U_1 \\ U_2 \\ U_3 \\ U_4 \\ U_5 \\ U_6 \end{array}\begin{array}{c} \begin{array}{cccccccccccc} J_1 & & J_2 & & J_3 & & J_4 & & J_5 & & J_6 & \end{array} \\ \begin{pmatrix} \{K_1, & K_2\} & \{K_1, & K_3\} & \{K_1, & K_4\} & \{K_2, & K_3\} & \{K_2, & K_4\} & \{K_3, & K_4\} \\ 0 & 0 & w^2_{1,1}=.58 & w^3_{3,1}=.42 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & w^4_{2,2}=.62 & w^4_{3,2}=.38 & 0 & 0 & 0 & 0 \\ w^1_{1,3}=.46 & w^1_{2,3}=.54 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w^5_{2,4}=.62 & w^5_{4,4}=.38 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w^6_{3,5}=.45 & w^6_{4,5}=.55 \\ 0 & 0 & 0 & 0 & w^3_{1,6}=.53 & w^3_{4,6}=.47 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \qquad (25)$$

TABLE 5. Weighted SC-user power in L2.

	J_1		J_2		J_3		J_4		J_5		J_6		Constraint
	$\{K_1,$	$K_2\}$	$\{K_1,$	$K_2\}$	$\{K_1,$	$K_2\}$	$\{K_1,$	$K_2\}$	$\{K_1,$	$K_2\}$	$\{K_1,$	$K_2\}$	C4: $\sum_{k\in\{S_{J_j}\}} P_{u,k} \leq P_{u,j}$
U_1	0	0	$P_{1,1}=1.16$	$P_{1,3}=.84$	0	0	0	0	0	0	0	0	2dB
U_2	0	0	0	0	0	0	$P_{2,2}=1.24$	$P_{2,3}=.76$	0	0	0	0	2dB
U_3	$P_{3,1}=.92$	$P_{3,2}=1.08$	0	0	0	0	0	0	0	0	0	0	2dB
U_4	0	0	0	0	0	0	0	0	$P_{4,2}=1.24$	$P_{4,4}=.76$	0	0	2dB
U_5	0	0	0	0	0	0	0	0	0	0	$P_{5,3}=.9$	$P_{5,4}=1.1$	2dB
U_6	0	0	0	0	$P_{6,1}=1.06$	$P_{6,4}=.94$	0	0	0	0	0	0	2dB
											C 3: $\sum_{j=1}^{J} P_{u,j} \leq P_{max}=$		12dB

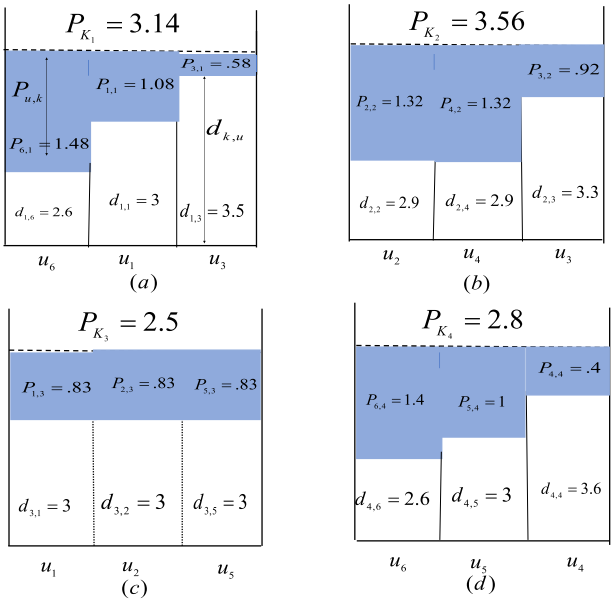


FIGURE 13. NOMA-SIC based GWF to estimate SC-user power in L4, (a)P_{K_1} = 3.14 dB (b)P_{K_2} = 3.56 dB (c)P_{K_3} = 2.5 dB (d)P_{K_4} = 2.8 dB.

L4: Using SC powers computed in L3, the SC-user powers in L4 can be computed by utilizing NOMA-SIC based GWF method. Fig. 13 shows the result of the SC-user power allocation, which performs the calculation of (20) and (21).

L5: Using SC-user power values in L4, the CB-user power is obtained as:

$$P_{U_1,J_2} = P_{U_1,K_1} + P_{U_1,K_3} = 1.08 + .83 = 1.91\text{dB}$$
$$P_{U_2,J_4} = P_{U_2,K_2} + P_{U_2,K_4} = 1.32 + .83 = 2.15\text{dB}$$
$$P_{U_3,J_1} = P_{U_3,K_1} + P_{U_3,K_2} = .58 + .92 = 1.5\text{dB}$$
$$P_{U_4,J_5} = P_{U_4,K_2} + P_{U_4,K_4} = 1.32 + .4 = 1.72\text{dB}$$
$$P_{U_5,J_6} = P_{U_5,K_3} + P_{U_5,K_4} = .83 + 1 = 1.83\text{dB}$$
$$P_{U_6,J_3} = P_{U_6,K_1} + P_{U_6,K_4} = 1.48 + 1.4 = 2.88\text{dB}$$
$$P_{U_1,J_2} + P_{U_2,J_4} + P_{U_3,J_1} + P_{U_4,J_5} + P_{U_5,J_6} + P_{U_6,J_3} = 12\text{dB}$$

REFERENCES

[1] D. Wubben, P. Rost, J. S. Bartelt, M. Lalam, V. Savin, M. Gorgoglione, A. Dekorsy, and G. Fettweis, "Benefits and impact of cloud computing on 5G signal processing: Flexible centralization through cloud-RAN," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 35–44, Nov. 2014.

[2] R. Wang, H. Hu, and X. Yang, "Potentials and challenges of C-RAN supporting multi-RATs toward 5G mobile networks," *IEEE Access*, vol. 2, pp. 1200–1208, 2014.

[3] A. Davydov, G. Morozov, I. Bolotin, and A. Papathanassiou, "Evaluation of joint transmission CoMP in C-RAN based LTE-A HetNets with large coordination areas," in *Proc. IEEE Globecom Workshops*, Dec. 2013, pp. 801–806.

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

[5] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann, "Cloud-RAN for mobile networks—A technology overview," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 405–426, 1st Quart., 2015.

[6] L. Ferdouse, A. Anpalagan, and S. Erkucuk, "Joint communication and computing resource allocation in 5G cloud radio access networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 9, pp. 9122–9135, Sep. 2019.

[7] A. Abdelnasser and E. Hossain, "Resource allocation for an OFDMA cloud-RAN of small cells underlaying a macrocell," *IEEE Trans. Mobile Comput.*, vol. 15, no. 11, pp. 2837–2850, Nov. 2016.

[8] T. Yunzheng, L. Long, L. Shang, and Z. Zhi, "A survey: Several technologies of non-orthogonal transmission for 5G," *China Commun.*, vol. 12, no. 10, pp. 1–15, Oct. 2015.

[9] M. Taherzadeh, H. Nikopour, A. Bayesteh, and H. Baligh, "SCMA codebook design," in *Proc. IEEE Veh. Technol. Conf.*, Sep. 2014, pp. 1–5.

[10] A. Ghaffari, M. Leonardon, Y. Savaria, C. Jego, and C. Leroux, "Improving performance of SCMA MPA decoders using estimation of conditional probabilities," in *Proc. IEEE Int. New Circuits Syst. Conf.*, Jun. 2017, pp. 21–24.

[11] X. Yue, Z. Qin, Y. Liu, S. Kang, and Y. Chen, "A unified framework for non-orthogonal multiple access," *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5346–5359, Nov. 2018.

[12] F. Zhou, Y. Wu, R. Q. Hu, Y. Wang, and K. K. Wong, "Energy-efficient NOMA enabled heterogeneous cloud radio access networks," *IEEE Netw.*, vol. 32, no. 2, pp. 152–160, Mar. 2018.

[13] M. Moltafet, N. Mokari, M. R. Javan, H. Saeedi, and H. Pishro-Nik, "A new multiple access technique for 5G: Power domain sparse code multiple access (PSMA)," *IEEE Access*, vol. 6, pp. 747–759, 2018.

[14] S. Sharma, K. Deka, V. Bhatia, and A. Gupta, "Joint power-domain and SCMA-based NOMA system for downlink in 5G and beyond," *IEEE Commun. Let.*, vol. 23, no. 6, pp. 971–974, Jun. 2019.

[15] W. Zhu, L. Qiu, and Z. Chen, "Joint subcarrier assignment and power allocation in downlink SCMA systems," in *Proc. IEEE Veh. Tech. Conf.*, Sep. 2017, pp. 1–5.

[16] M. A. Sedaghat and R. R. Müller, "On user pairing in uplink NOMA," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3474–3486, May 2018.

[17] B. Di, L. Song, and Y. Li, "Radio resource allocation for uplink sparse code multiple access SCMA networks using matching game," in *Proc. IEEE Int. Conf. Commun.*, May 2016, pp. 1–6.

[18] R. Ruby, S. Zhong, H. Yang, and K. Wu, "Enhanced uplink resource allocation in non-orthogonal multiple access systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1432–1444, Mar. 2018.

[19] F. Fang, H. Zhang, J. Cheng, and V. C. M. Leung, "Energy-efficient resource allocation for downlink non-orthogonal multiple access network," *IEEE Trans. Commun.*, vol. 64, no. 9, pp. 3722–3732, Sep. 2016.

[20] M. S. Ali, E. Hossain, A. Al-Dweik, and D. I. Kim, "Downlink power allocation for CoMP-NOMA in multi-cell networks," *IEEE Trans. Commun.*, vol. 66, no. 9, pp. 3982–3998, Sep. 2018.

[21] M. Moltafet, N. M. Yamchi, M. R. Javan, and P. Azmi, "Comparison study between PD-NOMA and SCMA," *IEEE Trans. Veh. Technol.*, vol. 67, no. 2, pp. 1830–1834, Feb. 2018.

[22] Z. Li, W. Chen, F. Wei, F. Wang, X. Xu, and Y. Chen, "Joint codebook assignment and power allocation for SCMA based on capacity with Gaussian input," in *Proc. IEEE Int. Conf. Commun. China*, Jul. 2016, pp. 1–6.

[23] J. Zhu, J. Wang, Y. Huang, S. He, X. You, and L. Yang, "On optimal power allocation for downlink non-orthogonal multiple access systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2744–2757, Dec. 2017.
[24] H. Zhang, B. Wang, C. Jiang, K. Long, A. Nallanathan, V. C. M. Leung, and H. V. Poor, "Energy efficient dynamic resource optimization in NOMA system," *IEEE Trans. Wireless Commun.*, vol. 17, no. 9, pp. 5671–5683, Sep. 2018.
[25] 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, Aug. 2009.
[26] L. Ferdouse, W. Ejaz, A. Anpalagan, and A. M. Khattak, "Joint workload scheduling and BBU allocation in Cloud-RAN for 5G networks," in *Proc. Symp Appl. Comput.* New York, NY, USA: ACM, 2017, pp. 621–627.
[27] M. Hasan and E. Hossain, "Resource allocation for network-integrated device-to-device communications using smart relays," in *Proc. IEEE Globecom Workshops*, Dec. 2013, pp. 591–596.
[28] B. Xu, Y. Chen, J. R. Carrión, and T. Zhang, "Resource allocation in energy-cooperation enabled two-tier NOMA HetNets toward green 5G," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2758–2770, Dec. 2017.
[29] Y. Li, M. Sheng, Z. Sun, Y. Sun, L. Liu, D. Zhai, and J. Li, "Cost-efficient codebook assignment and power allocation for energy efficiency maximization in SCMA networks," in *Proc. IEEE Vehic. Tech. Conf.*, Sep. 2016, pp. 1–5.
[30] P. He, L. Zhao, S. Zhou, and Z. Niu, "Water-filling: A geometric approach and its application to solve generalized radio resource allocation problems," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3637–3647, Jul. 2013.
[31] M. Salehi, H. Tabassum, and E. Hossain, "Accuracy of distance-based ranking of users in the analysis of NOMA systems," *IEEE Trans. Commun.*, vol. 67, no. 7, pp. 5069–5083, Jul. 2019.
[32] M. Basharat, M. Naeem, W. Ejaz, A. M. Khattak, A. Anpalagan, O. Alfandi, and H. S. Kim, "Non-orthogonal radio resource management for RF energy harvested 5G networks," *IEEE Access*, vol. 7, pp. 46550–46561, 2019.
[33] L. Ferdouse, A. Alnoman, A. Bulzacki, and A. Anpalagan, "Energy efficient multiple association in CoMP based 5G Cloud-RAN systems," in *Proc. IEEE Vehic. Tech. Conf.*, Sep. 2017, pp. 1–5.
[34] N. Wang, E. Hossain, and V. K. Bhargava, "Backhauling 5G small cells: A radio resource management perspective," *IEEE Wireless Commun.*, vol. 22, no. 5, pp. 41–49, Oct. 2015.

LILATUL FERDOUSE received the B.Sc. and M.S. degrees in computer engineering from Dhaka University, Dhaka, Bangladesh, in 2004 and 2006, respectively, the M.A.Sc. degree in electrical and computer engineering from Ryerson University, Toronto, ON, Canada, in 2015, and the Ph.D. degree from Ryerson University, in 2019. Her research interests are in resource allocation of next generation wireless communication systems, relay assisted cooperative networks, congestion, and overload control techniques in M2M, and the IoT networks.

SERHAT ERKUCUK (S'99–M'08–SM'19) received the B.Sc. degree in electrical engineering from Middle East Technical University, Ankara, Turkey, in 2001, the M.Sc. degree in electrical engineering from Ryerson University, Toronto, ON, Canada, in 2003, and the Ph.D. degree in engineering science from Simon Fraser University, Burnaby, BC, Canada, in 2007. He was an NSERC Postdoctoral Fellow at The University of British Columbia, Vancouver, Canada, in 2008. He then joined Kadir Has University, Istanbul, Turkey, where he is currently an Associate Professor. In 2018, he was a Visiting Professor with Ryerson University, where he conducted research on the design of small cells for 5G networks. His research interests are in physical layer design of emerging communication systems, wireless sensor networks, and communication theory. He is a Marie Curie Fellow and a recipient of the Governor's General Gold Medal.

ALAGAN ANPALAGAN (S'98–M'01–SM'04) received the B.A.Sc., M.A.Sc., and Ph.D. degrees from the University of Toronto, Canada, all in electrical engineering. He joined with the ELCE Department, Ryerson University, Canada, in 2001, and was promoted to Full Professor, in 2010. He served the department in administrative positions as the Associate Chair, the Program Director for electrical engineering, and the Graduate Program Director. During his sabbatical, he was a Visiting Professor with the Asian Institute of Technology, and Visiting Researcher at Kyoto University. His industrial experience includes working for three years with Bell Mobility, Nortel Networks, and IBM. He directs a research group working on radio resource management (RRM) and radio access and networking (RAN) areas within the WINCORE Laboratory. He has coauthored four edited books and two books in wireless communication and networking areas. He is a Registered Professional Engineer in the province of Ontario, Canada, and a Fellow of the Institution of Engineering and Technology (FIET) and a Fellow of the Engineering Institute of Canada (FEIC). He has served as the TPC Co-Chair, the IEEE VTC Fall 2017, the TPC Co-Chair, the IEEE INFOCOM'16: Workshop on Green and Sustainable Networking and Computing, and the IEEE Globecom15: SAC Green Communication and Computing, IEEE PIMRC'11: Cognitive Radio and Spectrum Management. He has served as the Vice Chair, the IEEE SIG on Green and Sustainable Networking and Computing with Cognition and Cooperation (2015–2018), the IEEE Canada Central Area Chair (2012–2014), the IEEE Toronto Section Chair (2006–2007), the ComSoc Toronto Chapter Chair (2004–2005), and the IEEE Canada Professional Activities Committee Chair (2009–2011). He was a recipient of the IEEE Canada J.M. Ham Outstanding Engineering Educator Award (2018), the YSGS Outstanding Contribution to Graduate Education Award (2017), the Deans Teaching Award (2011), Faculty Scholastic, Research and Creativity Award thrice from the Ryerson University. He was also a recipient of the IEEE M.B. Broughton Central Canada Service Award (2016), the Exemplary Editor Award from the IEEE ComSoc (2013), and the Editor-in-Chief Top10 Choice Award in Transactions on Emerging Telecommunications Technology (2012), and a coauthor of a paper that received the IEEE SPS Young Author Best Paper Award (2015). He has served as an Editor for the IEEE Communications Surveys and Tutorials (2012–2014), the IEEE Communications Letters (2010–2013), and the *EURASIP Journal of Wireless Communications and Networking* (2004–2009). He has also served as a Guest Editor for six special issues published in the IEEE, IET, and ACM.

ISAAC WOUNGANG received the Ph.D. degree in mathematics from the University of South, Toulon and Var, France, in 1994. From 1999 to 2002, he worked as a Senior Software Engineer at Nortel Networks, Ottawa, Canada. Since 2002, he has been with Ryerson University, where he is currently a Professor of computer science and the Director of the DABNEL Research Lab. His current research interests include radio resource management in next generation wireless networks, computer security, computational intelligence and machine learning applications, performance modeling, and optimization. He has published eight edited books, one authored books, and over 80 refereed journals and conference papers.