IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 8, 2021, accepted December 27, 2021, date of publication January 6, 2022, date of current version January 14, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3140814

Hybrid Controlled User Association and Resource Management for Energy-Efficient Green RANs With Limited Fronthaul

LI-HSIANG SHEN, (Member, IEEE), CHIA-LIN TSAI, CHIA-YU WANG, AND KAI-TEN FENG, (Senior Member, IEEE)

Department of Electrical and Computer Engineering, National Yang Ming Chiao Tung University, Hsinchu 300093, Taiwan

Corresponding author: Kai-Ten Feng (ktfeng@mail.nctu.edu.tw)

This work was supported in part by the Ministry of Science and Technology (MoST) under Grant 110-2221-E-A49-041-MY3, Grant 110-2224-E-A49-001, and Grant 110-2623-E-A49-001; in part by the Zyxel Communications Corporation, STEM Project; in part by the Higher Education Sprout Project of the National Yang Ming Chiao Tung University; and in part by the Ministry of Education (MOE), Taiwan.

ABSTRACT To alleviate the greenhouse effect, high network energy efficiency (EE) has become an important research target in wireless green communications. Therefore, an investigation for resource management to mitigate co-tier interference in a small-cell network (SCN) is provided. Moreover, with the merits of cloud radio access networks (C-RANs), small-cell base stations can be decomposed into central small cells (CSCs) and remote small cells (RSCs). To achieve coordination, split medium access control (MAC)-based functional splitting is adopted with a scheduler deployed at CSCs and retransmission functions implemented at RSCs. However, limited fronthaul has a compelling effect on RSCs owing to the requirements of user quality of service (QoS). In this paper, we propose a hybrid controlled user association and resource management (HARM) scheme to deal with the infeasibility of controlling all RSCs using a single CSC. A CSC performs the traffic-control-based user association and resource allocation (TURA) scheme for RSCs to mitigate intra-group interference within localized C-RANs, whereas the CSCs among separate C-RANs conduct cooperative resource competition (CRC) games to alleviate inter-group interference. Based on the regret-based learning algorithm, the proposed schemes are analytically proven to reach the correlated equilibrium (CE). The simulation results validate the capability of traffic control in the TURA scheme and the convergence of CRC. Moreover, a comparison of the proposed TURA, HARM, and CRC schemes with the benchmark is revealed. It is observed that the TURA scheme outperforms the other methods under ideal fronthaul control, whereas the HARM scheme can sustain EE performance considering practical implementation.

INDEX TERMS Cloud radio access networks, functional split, limited fronthaul, energy efficiency, green communication, user association, resource management, power allocation, game theory.

I. INTRODUCTION

As wireless traffic grows dramatically [1], fulfilling increasing user demand has become the dominant issue for next-generation wireless communication systems. The deployment of small-cell base stations (SBSs) has been promoted to resolve these concerns over recent years owing to their advantages of low transmit power and low cost [2], [3]. On the other hand, it is estimated that the energy consumption for information and communication technology (ICT) is rising at a rate of 15–20% per year. The ICT industry is responsible for 3% of worldwide annual electrical energy consumption, which gives rise to 2–4% of the world's carbon dioxide emissions and severe effects on the global environment [4]–[7]. Therefore, GreenTouch [8] suggests that not only improving the entire network's achievable capacity but also reducing the power consumption of base stations is significant to network energy efficiency (EE) for reducing the overall carbon footprint, which is the main focus in the state-of-the-art research [9]–[11]. Recently, the architecture of hyper-dense SBS deployment has been viewed as a key solution to satisfy the huge amount of traffic demand

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

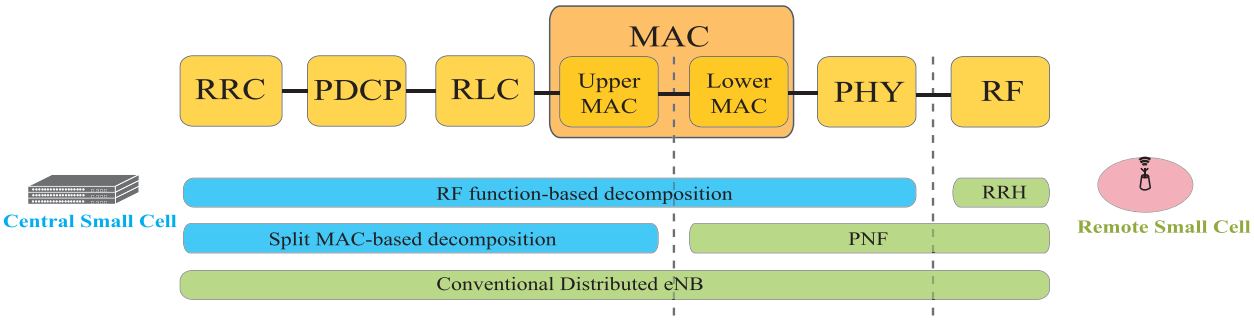


FIGURE 1. Functional split decompositions for a C-RAN.

while reducing energy consumption. Because small-cell networks (SCNs) shorten the distance between transmitters and receivers, the energy required for data transmission against pathloss can be reduced [12]. Owing to the merits of low transmit power and the tendency of dense deployment for high data rate requirements, the SCNs can be regarded as an energy-efficient wireless deployment for communication systems.

Furthermore, the emerging virtualization and cloud technologies shift the network functions from radio resource control (RRC), packet data convergence protocol (PDCP), radio link control (RLC), medium access control (MAC), physical (PHY) and radio frequency (RF) layers at the edge of SBSs to a central processing unit, which is called the cloud radio access network (C-RAN) [13]–[17]. The functional split capability in C-RANs has garnered attention from academia and industry [18]–[21][1]. The coordination among small cells in C-RANs can provide economic advantages and performance gains, including improved coordination, scalability enhancement, cost reduction, and more flexibility in network deployments [22]. The technologies of virtual network functions (VNFs) run aggregated small-cell functions from different base stations in virtual machines. Seven types of functional splits have been investigated [18] between the central small cells (CSCs) and remote small cells (RSCs) with the quality requirements of fronthaul, which connects CSCs to RSCs. The applicable network functions and hardware restrictions in different scenarios of functional splits are also discussed. In [23], the upper layer network functions of small cells were virtualized in a CSC to perform centralized management and coordination to serve its corresponding RCSs. On the other hand, the remaining functions reside in the edged small cells, namely RSCs, and autonomously perform lower-layer functionalities.

Fig. 1 shows the three types of functional splits extracted from [23], [24] that will be discussed in this paper. First, conventional distributed small cells as illustrated in the bottom case of Fig. 1 conduct network functions autonomously, and the existing link solutions can support fronthaul requirements with nationwide low-cost Internet Protocol (IP) networks. However, the limited capacity of non-ideal fronthaul imposes a constraint on the SBSs and leads to a performance bottleneck of achievable system throughput [23], [25]. Under practical consideration of fronthaul with limited capacity, there is a fundamental effect on user association while the fronthaul links are overloaded [26]. To guarantee the quality of service (QoS) of each user, the serving SBS with overloaded fronthaul will offload some users to other SBSs which is referred to as traffic control [27]. In [28], the authors aimed at maximizing the weighted-sum rate for fronthaul-constrained SCNs with carrier aggregation. Furthermore, it is crucial to appropriately allocate available frequency resources and transmit power of SBS to fulfill the user's QoS requirement. Non-cooperative games were adopted in [29]–[31] to enhance network capacity or EE by each SBS which selfishly determines the resource blocks (RBs) and power assignments based on its own utility functions, whereas the overall system performance will be degraded owing to a lack of coordination. Another proposed framework in [32] executes resource and power allocations by constructing a cooperative game between small cells for cross-tier and co-tier interference mitigation. Moreover, the scheme proposed in [33] achieves optimal network EE based on a cooperative game for subcarrier assignment. Research in [34] assigns subcarriers and allocates transmit power of SBS by evolutionary game theory, which analyzes the average interference between SBSs.

Another functional split virtualizes all the network functions in the CSC, whereas the RSCs only contain the radio-frequency (RF) processing unit for data transmission and reception (as shown in the top-most case of Fig. 1). This is the classic realization for the C-RAN architecture, and the VNFs are assigned to the most appropriate processor or hardware accelerator in CSCs to efficiently execute the corresponding network functions of base stations. Although higher system performance can be achieved owing to full coordination among small cells, the requirements of low latency and ideal fronthaul will result in considerable expenditure

[1]Note that the concept of network slicing is kind of different from the proposed functional split. Network slicing aims at partitioning different types of networks into several processing machines or planes [16], [17]. While, functional split is to separate network layer functions of RRC/PDCP/RLC/MAC/PHY/RF into either central, distributed, or front end units.

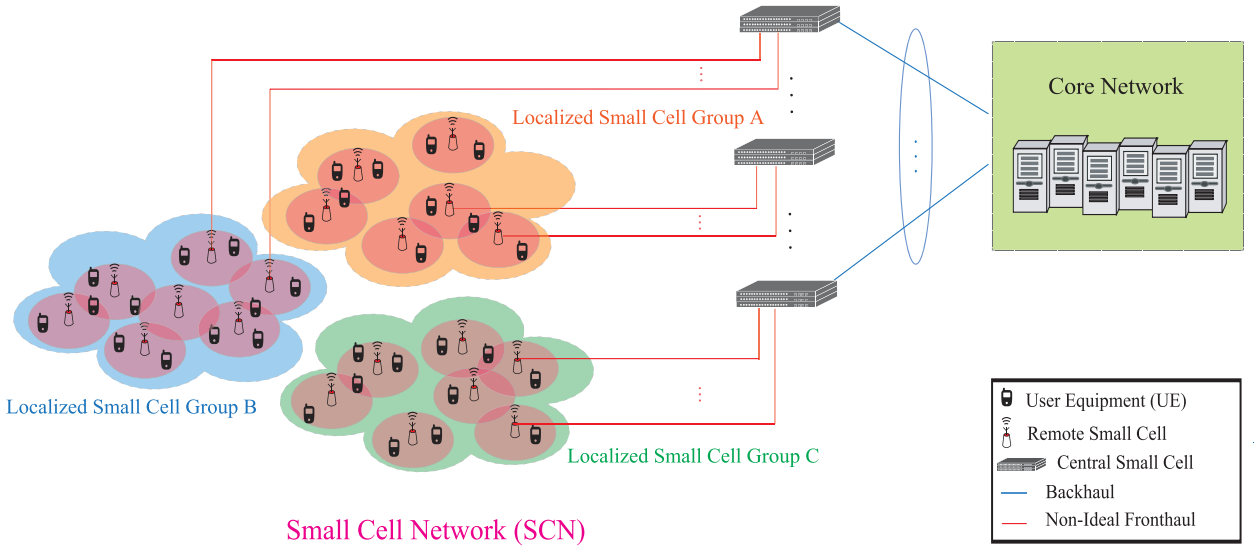


FIGURE 2. C-RAN network architecture for the proposed split-MAC-based SCN.

on network operators. Previous work in [35] conducted RB and power allocations for SCN by a centrally controlled unit, and an optimization problem was formulated to maximize the network EE. Nevertheless, the energy conservation and QoS requirements have not been considered in [28], [36]. The data rate requirements and restrictions of fronthaul capacity have been simultaneously considered in the literature [24], [28], [37]. In [24], an energy-efficient resource allocation algorithm for multi-cell orthogonal frequency division multiple access (OFDMA) systems was proposed. The work in [37] presented a joint resource allocation and admission control to minimize the sum of interference levels that macrocells can experience from small cells. Even though a near-optimal solution can be obtained, the signaling overhead and computational loadings are potential drawbacks to RF-based functional split management.

The above observations intuitively imply that there exists a tradeoff between the overall system performance and the deployment cost of fronthaul links. According to the analysis in [2], MAC can be divided into upper MAC as the scheduler and lower MAC as the hybrid automatic repeat request (HARQ) mechanism. In addition, the functional split of MAC can deliver the benefits of centralization, but only requires a small increase in transporting data in such a network scenario. This is well aligned with the existing multivendor ecosystem for telecom operators based on the functional application platform interface. However, it is apparently infeasible for a CSC to control a large number of RSCs in realistic communication systems owing to hardware limitations. As a result, a practical scale for the implementation of dense SCNs is analyzed in this paper, where a CSC will be in charge of radio resource management (RRM) for RSCs in localized small-cell groups such as shopping malls or commercial buildings under split-MAC-based network functions. Interference management with limited fronthaul capacity will be invoked within a localized small-cell group. Furthermore, a scheme for interference mitigation among localized small-cell groups is proposed in this paper. With the preponderance of split-MAC-based functional splitting, this paper proposes a framework for subchannel and RSC transmit power allocation to maximize EE centrally by a CSC in its localized serving small-cell group. The traffic control and small-cell on/off mechanisms were also designed considering the limited capacity for non-ideal fronthauls. Traffic control occurs with overloaded fronthaul of serving RSC, and this overloaded RSC offloads the users to one of the nearby RSCs, which retain sufficient fronthaul capacity to serve the user. Meanwhile, the RSC will also tend to offload the associated users to others when its loading is low, so as to turn off the RSC for energy saving. Hence, not only the required QoS under the circumstance of limited fronthaul capacity, but also the power conservation of RSCs can be achieved by the mechanism of user association. The contributions of this paper are as follows.

- We conceive a joint EE optimization problem for the traffic-control-based user association, transmit power and RB allocation problem constrained by limited fronthaul capacity, user QoS and allowable power control. To the best of our knowledge, the merits of this joint optimization can not only reduce the computational load of CSCs but also mitigate the power consumption to enhance the network performance. Unlike ideal implementation of VNFs co-located within a CSC, the interference alleviation for edge users among all localized small cells imposes a challenge to coordinate RRM with limited and asynchronous exchanged information, which is not comprehensively considered in existing studies.

- We propose a hybrid controlled user association and resource management (HARM) scheme, which consists of distributed/centralized RRM for split-MAC-based functional splitting SCNs in C-RANs restricted by fronthaul capacity. A centralized RRM of traffic-control-based user association and resource allocation (TURA) is conducted in a localized C-RAN, whereas a distributed learning manner of cooperative resource competition (CRC) is designed for subchannel and transmit power allocation among different CSCs, which is analytically proven as a stable network. The proposed scheme improves the network EE by effectively managing inter- and intra-group interferences, which dynamically adjusts the policy according to the power competition among CSCs.
- The performance results have demonstrated the effectiveness of proposed HARM by verifying its convergence of CRC. Additionally, EE of the proposed TURA scheme is better than the existing methods with respect to different numbers of users, fronthaul capacity limitations, allowable signaling power, and asynchronous error ratios. This is because the TURA scheme can conduct fully-centralized management under the consideration of an ideal fronthaul. Despite the slightly lower EE performance of HARM than the ideal TURA, HARM can sustain an appropriate EE compared with existing methods under a feasible fronthaul implementation of practical green RANs.

The remainder of this paper is organized as follows. Detailed descriptions of the system model and problem formulation of the proposed HARM algorithm are provided in Section II. Section III illustrates the proposed TURA scheme for the SCN in localized C-RANs. Section IV formulates the CRC scheme among multiple localized C-RANs and adopts a distributed algorithm to reach a correlated equilibrium (CE). The performance of the proposed framework is evaluated in Section V. Finally, Section VI concludes the paper.

Notation: A bold capital letter denotes a matrix and $|\cdot|$ as the absolute value. The operator $\max(\cdot)$ returns the largest value in an array. We use $\mathbb{1}(A)$ to denote an indicator function equal to 1 when event A occurs and 0 otherwise. A random variable with a value between a and b is generated by $rand(a, b)$.

II. SYSTEM MODEL AND PROBLEM FORMULATION

Detailed descriptions of the architecture and operating process of the proposed split-MAC-based SCN are provided in the first and second subsections. In addition, the energy-efficient optimization problem of resource allocation with traffic control and small-cell on/off mechanisms under relevant constraints of subchannels, RSC transmit power, and capacity of fronthaul is formulated in the third subsection.

A. SYSTEM MODEL

Fig. 2 illustrates a downlink OFDMA cellular network, which is divided into $\mathcal{C} = \{1, \ldots, C\}$ localized small-cell groups. There exists a set of $\mathcal{S} = \{1, \ldots, S\}$ open-access RSCs and a set of $\mathcal{K} = \{1, \ldots, K\}$ serving user equipment (UE) in the SCN. Each RSC connects to its serving CSC by

TABLE 1. Definition of system parameters.

Parameters	Notation
Set of C-RAN, RSC, UE, channel	$\mathcal{C}, \mathcal{S}, \mathcal{K}, \mathcal{N}$
Number of C-RANs, RSCs, UEs, channels	C, S, K, N
Received SINR	$\gamma^n_{c,s,k}$
Channel gain	$g^n_{c,s,k}$
Interference power	$I^n_{c,s,k}$
Operating bandwidth	W
Noise power spectral density	N_0
Individual capacity	$R^n_{c,s,k}$
Sum rate within a localized C-RAN	R_c
Individual/Total power of signaling	$P^{(\mathrm{O})}, P^{(\mathrm{TC})}$
Circuit power in active, sleep mode	$P^{(\mathrm{CA})}, P^{(\mathrm{CS})}$
Total power consumption	P_c
Individual energy efficiency	η_c
Initial state of user association	$\bar{\phi}_{c,s,k}$
Policy of user association	$\mathbf{\Phi}_c = \{\phi_{c,s,k}\}$
Policy of subchannel allocation	$\mathbf{\Psi}_c = \{\psi_{c,s,k}\}$
Policy of transmit power allocation	$\mathbf{P}_c = \{p^n_{c,s,k}\}$
Upper bound of transmit power of CRC	$\bar{\mathbf{P}}_c = \{\bar{P}^n_c\}$
The bounded C-RAN interference of CRC	$\kappa^{n,\ell}_{c,s,k}$
Total network EE	$\eta(\mathbf{\Phi}, \mathbf{\Psi}, \mathbf{P})$
Maximum transmit power of each RSC	$P^{(\max)}_{c,s}$
Minimum UE QoS constraint	$R^{(\min)}_{c,k}$
Maximum fronthaul capacity constraint	$B^{(\max)}_{c,s}$
EE of localized C-RAN	$\tilde{\eta}_c(\mathbf{\Phi}_c, \mathbf{\Psi}_c, \mathbf{P}_c)$
Rate of localized C-RAN	$\tilde{R}_c(\mathbf{\Phi}_c, \mathbf{\Psi}_c, \mathbf{P}_c)$
Estimated total interference	$\tilde{I}^n_{c,s,k}$
Erroneous power from other C-RANs	$\tilde{p}^n_{t,k,i}$
Fitness function of QPSO in TURA	$F(\mathbf{\Phi}_c, \mathbf{\Psi}_c, \mathbf{P}_c)$
Penalty factor	α
Penalty function	$\Delta(\mathbf{\Phi}_c, \mathbf{\Psi}_c, \mathbf{P}_c)$
Total/Candidate decision TURA policy	$\mathbf{X}_c, \mathbf{X}_i(t)$
Threshold of fitness, iteration	$F^{(\mathrm{th})}, T$
Best history of QPSO particle	$\mathbf{X}^{(\mathrm{HB})}_i(t)$
Globally best solution among all particles	$\mathbf{X}^{(\mathrm{GB})}(t)$
Weighted mean of elite solutions	$\mathbf{X}^{(\mathrm{Mean})}(t)$
Attractor solution between local/global	$\mathbf{X}^{(\mathrm{SP})}_i(t)$
Parameters of QPSO in TURA	$\beta(t), \mu_i(t), \varepsilon_i(t)$
Significance of local/global solution	$\lambda_i(t)$
Threshold of particle update	δ
Convergence parameters of TURA	$\beta^{(\max)}, \beta^{(\min)}$
Fitness gap ratio	$\tau_i(t)$
CRC cooperative game	$\mathcal{G}$
Joint allocation indicator	$\kappa^n_{c,s,k}$
Power-level based joint allocation indicator	$\kappa^{n,\ell}_{c,s,k}$
Number of power levels	L_C
Utility function	U_c
Total action set	$\mathcal{A}$
Total/Selected action set per CSC	$\mathcal{A}_c, \mathbf{a}_c$
Total/Selected off-policy set per CSC	$\mathcal{A}_{-c}, \mathbf{A}_{-c}$
Unused action policy	$\hat{\mathbf{a}}_c$
Average regret value	$R^t_c(\mathbf{a}_c, \hat{\mathbf{a}}_c)$
Utility difference	$D^t_c(\mathbf{a}_c, \hat{\mathbf{a}}_c)$
CRC policy update/exchange probability	$\mathbf{w}^t_c(\hat{\mathbf{a}}^t_c), \tilde{\mathbf{w}}^t_{-c}(\hat{\mathbf{a}}^t_{-c})$
Normalization variable of CRC	ξ
Asynchronous error parameters	$\rho, \Delta h$
Threshold of CRC game	$\theta^{(\mathrm{th})}$
Performance of QoS outage	Δ_{out}

non-ideal fronthaul links, which possess the restriction of limited capacity in the rest of this paper within a localized small-cell group. The CSCs in different small-cell groups communicate with the core network by backhaul links. In the proposed split-MAC-based SCN, each CSC will conduct centralized resource allocation for its serving RSCs in a localized small-cell group. In other words, a localized small-cell group can be regarded as a localized C-RAN. It is considered that the channel state information at the transmitter (CSIT) can be measured by UEs and feedback to CSC through the RSCs. Furthermore, the channel set is defined as $\mathcal{N} = \{1, \ldots, N\}$ which has N available subchannels in the system. All the RSCs share the entire frequency band, which leads to inter-cell interference between the RSCs and a subchannel can only be assigned to a single UE. As a result, the signal-to-interference-plus-noise ratio (SINR) $\gamma^n_{c,s,k}$ of UE k served by RSC s in localized C-RAN c on the subchannel n is given as

$$\gamma^n_{c,s,k} = \frac{p^n_{c,s,k} g^n_{c,s,k}}{I^n_{c,s,k} + N_0 W}, \tag{1}$$

where $g^n_{c,s,k}$ and $p^n_{c,s,k}$ are the deterministic channel gain and transmit power from RSC s in the localized C-RAN c to UE k on subchannel n. Here N_0 is the power spectral density of additive white Gaussian noise (AWGN), and W is the bandwidth of a subchannel. The term $I^n_{c,s,k}$ in the denominator of (1) is represented in (2), which includes both the intra-group interference in a localized C-RAN (i.e., the first two terms) and the co-channel inter-group interference from other localized C-RANs to the localized C-RAN c (i.e., the third term):

$$\begin{aligned} I^n_{c,s,k} = & \sum_{\substack{j=1 \\ j\neq k}}^{K} \phi_{c,s,j} \psi^n_{c,s,j} p^n_{c,s,j} g^n_{c,s,k} \\ & + \sum_{\substack{i=1 \\ i\neq s}}^{S} \sum_{\substack{j=1 \\ j\neq k}}^{K} \phi_{c,i,j} \psi^n_{c,i,j} p^n_{c,i,j} g^n_{c,i,k} \\ & + \sum_{\substack{t=1 \\ t\neq c}}^{C} \sum_{\substack{i=1 \\ i\neq s}}^{S} \sum_{\substack{j=1 \\ j\neq k}}^{K} \phi_{t,i,j} \psi^n_{t,i,j} p^n_{t,i,j} g^n_{t,i,k}, \end{aligned} \tag{2}$$

where $\phi_{t,i,j} \in \{0, 1\}$ indicates whether user j is associated with RSC i in localized C-RAN t, and $\psi^n_{t,i,j} \in \{0, 1\}$ is the assignment of subchannel n to user j served by RSC i in localized C-RAN t. We notice that all serving UEs are in the connected mode; however, they may not be assigned resource when encountering limited resources or worse channel conditions. Given the SINR calculated in (1), the achievable data rate $R^n_{c,s,k}$ for user k served by RSC s in the localized C-RAN c on the subchannel n can be formulated based on the Shannon capacity as

$$R^n_{c,s,k} = W \log_2 \left(1 + \gamma^n_{c,s,k}\right). \tag{3}$$

Therefore, the sum rate R_c within a localized C-RAN c is acquired as

$$R_c = \sum_{s=1}^{S} \sum_{k=1}^{K} \phi_{c,s,k} \sum_{n=1}^{N} \psi^n_{c,s,k} R^n_{c,s,k}. \tag{4}$$

Furthermore, the transmit power of the RSC s in the localized C-RAN c is represented as

$$P^{(Tx)}_{c,s} = \sum_{k=1}^{K} \phi_{c,s,k} \sum_{n=1}^{N} \psi^n_{c,s,k} p^n_{c,s,k}. \tag{5}$$

The users are initially connected to RSCs with the largest reference signal receiving power (RSRP) and may be offloaded to other cells based on the conditions of traffic load and available fronthaul capacity. A signal power overhead $P^{(O)}$ is considered to reduce the ping-pong effect, which indicates that handovers back-and-forth between two RSCs contribute to system overloading. Therefore, the power overhead in RSC s resulting from traffic control is formulated as

$$P^{(TC)}_{c,s} = P^{(O)} \sum_{k=1}^{K} \phi_{c,s,k} \mathbb{1}\left(|\phi_{c,s,k} - \bar{\phi}_{c,s,k}| > 0\right), \tag{6}$$

where $\bar{\phi}_{c,s,k}$ represents the initial state of user association. Based on (5) and (6), the total power consumption P_c within a localized C-RAN can be modeled as

$$\begin{aligned} P_c = & \sum_{s=1}^{S} \mathbb{1}\left(\sum_{k=1}^{K} \phi_{c,s,k} > 0\right) \left(P^{(Tx)}_{c,s} + P^{(TC)}_{c,s} + P^{(CA)}\right) \\ & + \sum_{s=1}^{S} \left(1 - \mathbb{1}\left(\sum_{k=1}^{K} \phi_{c,s,k} > 0\right)\right) P^{(CS)}, \end{aligned} \tag{7}$$

where $P^{(CA)}$ and $P^{(CS)}$ are the circuit power consumption of the RSC in active mode and sleep mode, respectively. As switching the RSC with low traffic load into sleep mode is an efficient approach to reduce the power consumption of SBSs for green communication [38], $P^{(CA)}$ and $P^{(CS)}$ are consequently taken into consideration in the power model. The indicator function $\mathbb{1}\left(\sum_{k=1}^{K} \phi_{c,s,k} > 0\right)$ can be viewed as an RSC on/off strategy that is equal to 0 when there is no user associated with the RSC, and the RSC will enter sleep mode for power saving. The indicator function becomes 1 when there are a single or multiple users associated with the RSC, and the RSC is in active mode. Accordingly, the EE of a localized C-RAN, which is defined as the ratio of total achievable data rate to the total power consumption of SBSs, can be expressed from (4) and (7) as

$$\eta_c = \frac{R_c}{P_c}. \tag{8}$$

B. OPERATIONAL PROCESS FOR THE PROPOSED HARM SCHEME

The operating flow chart of the proposed HARM system for mitigating both inter- and intra-group interference is shown in Fig. 3. Because the interference from the UEs can be

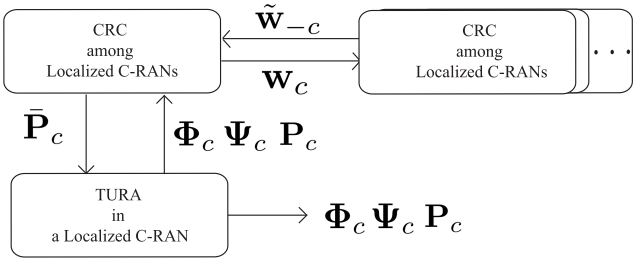


FIGURE 3. Operating flow chart for interference management of the proposed HARM scheme.

decomposed into intra-group interference within the same localized C-RAN and inter-group interference from different localized C-RANs, the HARM scheme is proposed to overcome the above-mentioned problems by repeatedly implementing the CRC and TURA algorithms. By implementing the proposed TURA algorithm, the users will be associated with their serving RSC and allocated with the proper configuration of subchannel and RSC transmit power considering the limited capacity of fronthaul and QoS requirements. Intra-group interference can be mitigated by the central control of CSCs within each localized C-RAN. The decision strategies in the TURA scheme, including user association, subchannel allocation, and transmit power allocation within a localized C-RAN are defined as $\Phi_c = \{\phi_{c,s,k} | 1 \le s \le S, 1 \le k \le K\}$, $\Psi_c = \{\psi^n_{c,s,k} | 1 \le s \le S, 1 \le k \le K, 1 \le n \le N\}$, and $\mathbf{P}_c = \{p^n_{c,s,k} | 1 \le s \le S, 1 \le k \le K, 1 \le n \le N\}$, respectively.

Furthermore, a CRC game is performed between CSCs based on the EE of their own localized C-RAN to alleviate inter-group interference. The decision strategies Φ_c, Ψ_c, and $\mathbf{P}_c$ obtained in the TURA scheme will be utilized by the localized CSC to determine the probability of strategy for resource allocation $\mathbf{w}_c$, and will be delivered from a localized C-RAN to other C-RANs. It is considered that erroneous information $\tilde{\mathbf{w}}_{-c}$ will be received by a CSC from other localized C-RANs owing to asynchronous communication between CSCs. The detailed descriptions of $\mathbf{w}_c$ and $\tilde{\mathbf{w}}_{-c}$ are given in Subsection IV-C. Moreover, with the adoption of the proposed CRC scheme, the set of upper bounds for transmit power on each subchannel in a localized C-RAN c can be obtained as $\bar{\mathbf{P}}_c = \{\bar{P}^n_c | 1 \le n \le N\}$, where the total transmit power $\bar{P}^n_c$ on subchannel n is $\bar{P}^n_c = \sum_{s=1}^{S} \sum_{k=1}^{K} \sum_{\ell=1}^{L_s} \kappa^{n,\ell}_{c,s,k}, \kappa^{n,\ell}_{c,s,k}$ is the constrained inter-group interference determined by proposed CRC scheme (as shown in Subsection IV-A), and L_s is the total power level. In other words, under the circumstance of bounded inter-group interference, each CSC executes the TURA scheme to determine user association, RSC transmit power allocation, and subchannel assignment to mitigate intra-group interference. We can then effectively provide sum-rate enhancement and reduce power consumption under limited fronthaul capacity within a localized C-RAN. As a result, the proposed HARM scheme is performed by repeatedly executing the TURA and CRC algorithms until it converges.

C. PROBLEM FORMULATION

The objective of this work is to maximize EE through subchannel assignment, power allocation, user association, and small-cell on/off mechanisms under the constraints of the data-rate requirement of each UE, maximum transmit power allowance of each RSC, and limited fronthaul capacity. The optimization problem of resource allocation can be formulated to acquire the transmission policies for Φ_c, Ψ_c, and $\mathbf{P}_c$ to improve the network EE, which is stated as follows:

$$\max_{\Phi,\Psi,\mathbf{P}} \ \eta(\Phi, \Psi, \mathbf{P}) = \sum_{c=1}^{C} \eta_c(\Phi_c, \Psi_c, \mathbf{P}_c) \tag{9a}$$

$$\text{s.t} \quad P^{(\mathrm{Tx})}_{c,s} \le P^{(\max)}_{c,s}, \quad \forall c, \ \forall s, \tag{9b}$$

$$\sum_{s=1}^{S} \phi_{c,s,k} \sum_{n=1}^{N} \psi^n_{c,s,k} R^n_{c,s,k} \ge R^{(\min)}_{c,k}, \quad \forall c, \forall k, \tag{9c}$$

$$\sum_{k=1}^{K} \phi_{c,s,k} \sum_{n=1}^{N} \psi^n_{c,s,k} R^n_{c,s,k} \le B^{(\max)}_{c,s}, \quad \forall c, \forall s, \tag{9d}$$

$$\sum_{s=1}^{S} \phi_{c,s,k} = 1, \quad \forall c, \ \forall k, \tag{9e}$$

$$\phi_{c,s,k}, \psi^n_{c,s,k} \in \{0, 1\}, \quad \forall c, \ \forall s, \ \forall k, \ \forall n, \tag{9f}$$

$$p^n_{c,s,k} \ge 0, \quad \forall c, \ \forall s, \ \forall k, \ \forall n. \tag{9g}$$

The parameter Φ in (9a) is the set of user association configurations, and $\Phi = \{\Phi_c \in |1 \le c \le C\}$. The sets of decision policies for subchannel assignment and RSC power allocation are defined as $\Psi = \{\Psi_c \in |1 \le c \le C\}$ and $\mathbf{P} = \{\mathbf{P}_c \in |1 \le c \le C\}$. In (9b), $P^{(\max)}_{c,s}$ is the maximum transmit power of each RSC, which restricts the sum of the allocated power on all subchannels. Constraint (9c) specifies that each user achieves its target data rate $R^{(\min)}_{c,k}$ according to the QoS requirement. Constraint (9d) shows that the sum rate of each RSC should be less than the allowance of the fronthaul capacity $B^{(\max)}_{c,s}$. Furthermore, (9e) states that each user can be served by only a single RSC under a localized C-RAN. The constraint in (9f) indicates that both $\phi_{c,s,k}$ and $\psi^n_{c,s,k}$ are binary integer variables for user association and subchannel assignment, respectively. Constraint (9g) defines the power-allocation parameters as non-negative values.

III. PROPOSED TURA SCHEME WITHIN A LOCALIZED C-RAN

In this section, our proposed TURA scheme is presented, which centrally performs resource allocation by the CSC for its corresponding RSCs in the localized C-RAN. The TURA scheme can mitigate intra-group interference and provide higher network capacity by adopting efficient traffic offloading, user association, and resource allocation. Moreover, proper configuration of user association can not only

satisfy the QoS requirement under limited fronthaul capacity, but can also turn off the lightly loaded RSC to conserve the transmit power, which gives rise to higher EE. Based on (9), the optimization problem for the proposed TURA algorithm can be formulated for a localized C-RAN c as follows:

$$\max_{\Phi_c,\Psi_c,\mathbf{P}_c} \tilde{\eta}_c(\Phi_c, \Psi_c, \mathbf{P}_c) \tag{10a}$$

$$\text{s.t (9b)–(9g)}, \tag{10b}$$

$$\sum_{s=1}^{S}\sum_{k=1}^{K} \phi_{c,s,k}\psi^n_{c,s,k}p^n_{c,s,k} \le \bar{P}^n_c, \quad \forall c, \forall n. \tag{10c}$$

Note that additional constraint (10c) is introduced which indicates that the assigned power on each subchannel n in a localized C-RAN c is constrained by the available power $\bar{P}^n_c$. The objective function $\tilde{\eta}_c$ in (10a) is the EE for localized C-RAN c based on the estimated SINR, which is calculated as

$$\tilde{\eta}_c = \frac{\tilde{R}_c(\Phi_c, \Psi_c, \mathbf{P}_c)}{P_c(\Phi_c, \Psi_c, \mathbf{P}_c)}, \tag{11}$$

where the estimated sum-rate $\tilde{R}_c$ in (11) for the localized C-RAN c is obtained as

$$\tilde{R}_c = \sum_{s=1}^{S}\sum_{k=1}^{K} \phi_{c,s,k} \sum_{n=1}^{N} \psi^n_{c,s,k} W \log_2\left(1 + \frac{p^n_{c,s,k}g^n_{c,s,k}}{\tilde{I}_{c,s,k} + N_0 W}\right). \tag{12}$$

The term $\tilde{I}_{c,s,k}$ in (12) is the estimated interference consisting of intra-group and erroneous inter-group interferences owing to asynchronous information exchanged between localized C-RANs, which can be acquired as

$$\begin{aligned}\tilde{I}^n_{c,s,k} = &\sum_{\substack{j=1\\ j\neq k}}^{K} \phi_{c,s,i}\psi^n_{c,s,j}p^n_{c,s,j}g^n_{c,s,k} \\ &+ \sum_{\substack{j=1\\ j\neq s}}^{S}\sum_{\substack{i=1\\ i\neq k}}^{K} \phi_{c,j,i}\psi^n_{c,j,i}p^n_{c,j,i}g^n_{c,j,k} \\ &+ \sum_{\substack{t=1\\ t\neq c}}^{C}\sum_{\substack{j=1\\ j\neq s}}^{S}\sum_{\substack{i=1\\ i\neq k}}^{K} \phi_{t,j,i}\psi^n_{t,j,i}\tilde{p}^n_{t,j,i}g^n_{t,j,k},\end{aligned} \tag{13}$$

where $\tilde{p}^n_{t,j,i}$ indicates the erroneous information of the transmit power from the other localized C-RANs. Because the proposed optimization problem of resource allocation in (10) contains a nonlinear objective function, which is the ratio of three transmit decision policies Φ_c, Ψ_c, and $\mathbf{P}_c$, it is difficult to solve this problem using conventional linear programming methods. In addition, the optimization problem is non-convex with respect to Φ_c, Ψ_c, and $\mathbf{P}_c$ because the discrete variables and co-channel interference are also considered. In general, the problem can be solved by an exhaustive search method, which tries all the allocated configurations for user association, subchannels, and transmit power of the RSC. However, the computational complexity increases exponentially with the number of RSCs. For conventional optimization methods to deal with this non-convex optimization problem, the discrete variables must be relaxed to continuous variables by adopting programming techniques and transforming the original problem into a convex problem. Nevertheless, the transformed optimization problem cannot be solved directly to obtain near-optimal solutions [39], [40].

To overcome the difficulty of dealing with the optimization problem in (10), the TURA algorithm with stochastic processes is proposed and described in this subsection. Particle swarm optimization (PSO), motivated by the social behavior of bird flocks or fish schools, has gained increasing popularity during the last decade owing to its effectiveness in conducting difficult optimization tasks, especially in resource allocation of wireless systems. The potential solutions of the resource allocation problem are called particles in PSO, and the particles will spread through the problem space to achieve the final resource configuration by considering historical data and the current best particle. Compared with the well-known genetic algorithm (GA), the main merits of PSO over GA rely on the momentum effect of velocity vectors for particle movement, which can quickly move the current best solution of each candidate particle to the global best solution, resulting in faster algorithm convergence. However, the solutions from PSO can easily be trapped in a local optimum, which can be effectively alleviated by adopting quantum-behaved particle swarm optimization (QPSO) [41], [42]. A random number generator is utilized in QPSO with a certain probability distribution to simulate the particle trajectories to provide the global convergence of particles. Furthermore, unlike PSO, QPSO does not require velocity vectors for particles and also possesses fewer parameters to adjust, which makes it easier to implement in realistic wireless systems. The position of a particle in QPSO, that is, the candidate solution $\{\Phi_c, \Psi_c, \mathbf{P}_c\}$ of (10), can be iteratively updated based on the particle fitness and evolution process to approach the optimal solution.

A. FITNESS FUNCTION AND TRANSFORMATION FOR UNCONSTRAINED FORM

In the proposed TURA scheme, each potential solution becomes a candidate by evaluating the quality of the fitness function. However, the fitness function is generally in an unconstrained form, and a transformation from the constrained objective function in (10) is required. To tackle this difficulty, the penalty function is adopted to transform the original optimization problem in (10) into an unconstrained problem [41], where the fitness function is defined as the gain between the reward and penalty functions as

$$F(\Phi_c, \Psi_c, \mathbf{P}_c) = \tilde{\eta}_c(\Phi_c, \Psi_c, \mathbf{P}_c) - \alpha\Delta(\Phi_c, \Psi_c, \mathbf{P}_c), \tag{14}$$

where α is the penalty factor, which is a parameter for the particle to balance the fitness function between the EE performance and the penalty that does not satisfy the constraints. Moreover, $\Delta(\Phi_c, \Psi_c, \mathbf{P}_c)$ in (14) represents the penalty function that can be obtained as (15), as shown at the bottom of the next page. It can be observed from (15) that the

value of the penalty function relies on the gaps between the transmission policies and the corresponding constraints. The more constraints that are unsatisfied, the higher penalty the particle needs to suffer from conducting this decision of transmission policies, which affects the direction of the path to either become the globally best position or is eliminated when compared with its historical particles during the evolution process.

B. OPERATION PROCESS FOR THE TURA SCHEME

Because the updating process for each variable is implemented in the same manner, let $\mathbf{X}_c$ generalize the decision policies $\{\Phi_c, \Psi_c, \mathbf{P}_c\}$. In each iteration t, there are I candidate solutions of the optimization problem in (10) to be chosen from the search space for all possible solution sets. The detailed descriptions of the TURA scheme are as follows.

1) INITIALIZATION

All I candidate solutions were initialized at the beginning. For notational simplicity, the ith candidate solution at iteration t is denoted by $\mathbf{X}_i(t)$. The fitness function in (14) of each candidate solution is calculated, and the threshold of the fitness function $F^{(th)}$ for the algorithm is determined to achieve the convergent solution.

2) EVOLUTION

After the initialization of I particles, we can obtain the best solution of the candidate ith particle in iteration t, $\mathbf{X}_i^{(HB)}(t)$, which contributes to the largest value of the fitness function in (14) in its history. Moreover, the global best solution among all I particles at iteration t, which is denoted as $\mathbf{X}^{(GB)}(t)$, can also be acquired. The weighted mean of the I elite candidate solutions at iteration t can be calculated as

$$\mathbf{X}^{(Mean)}(t) = \frac{1}{I}\sum_{i=1}^{I}\mathbf{X}_i^{(HB)}(t). \tag{16}$$

Consequently, the ith candidate solution in iteration t is updated by the evolution equation [43], [44] in (17), as shown at the bottom of the page, where $\mu_i(t) = rand(0,1)$ and $\varepsilon_i(t) = rand(0,1)$, along with a given threshold of δ. The term $\mathbf{X}_i^{(SP)}(t)$ in (17) is the attractor between local and global optimal solutions for candidate i in iteration t, which is given by

$$\mathbf{X}_i^{(SP)}(t) = \lambda_i(t)\cdot\mathbf{X}_i^{(HB)}(t) + (1-\lambda_i(t))\cdot\mathbf{X}^{(GB)}(t), \tag{18}$$

where $\lambda_i(t) = rand(0,1)$. Based on the evolution equation in (17), the candidate solutions can be obtained based on $\mathbf{X}_i^{(SP)}(t)$, $\mathbf{X}_i^{(Mean)}(t)$, and $\mathbf{X}_i(t)$. The new starting point $\mathbf{X}_i^{(SP)}(t)$ for the candidate solution at the next iteration $t+1$ depends on the best solution in the history of the ith candidate and the global best solution among all the I candidates at iteration t. The parameter $\lambda_i(t)$ in (18) is a random variable for the candidate at the next iteration that starts from a new position, which cross-correlates the historically best solution with the global best solution instead of starting from either $\mathbf{X}_i^{(HB)}(t)$ or $\mathbf{X}^{(GB)}(t)$ unilaterally. Moreover, the second term $\beta(t)|\mathbf{X}^{(Mean)}(t) - \mathbf{X}_i(t)|\cdot\ln\left(\frac{1}{\mu_i(t)}\right)$ of (17) affects the speed at which the position of the next ith candidate solution converges to the final solution. The term $\beta(t)$ is a coefficient that influences the convergence speed of the algorithm [1], which is given by

$$\beta(t) = \left(\beta^{(max)} - \beta^{(min)}\right)\cdot\frac{T-t}{T} + \beta^{(min)}, \tag{19}$$

where $\beta^{(max)}$ and $\beta^{(min)}$ are the maximum and minimum search ranges in the solution space, respectively. The maximum number of iterations is denoted as T. It can be observed from (19) that $\beta(t)$ decreases linearly with t because $\mathbf{X}_i(t)$ approaches convergence for large t. The difference between $\mathbf{X}^{(Mean)}(t)$ and $\mathbf{X}_i(t)$ also affects the speed of convergence. If the difference is large, which means that this candidate

$$\begin{aligned}\Delta(\Phi_c,\Psi_c,\mathbf{P}_c) = &\sum_{s=1}^{S}\left[\max\left(0, P_{c,s}^{(Tx)} - P_{c,s}^{(max)}\right)\right]^2 + \sum_{s=1}^{S}\left[\max\left(0, \sum_{k=1}^{K}\phi_{c,s,k}\sum_{n=1}^{N}\psi_{c,s,k}^{n}R_{c,s,k}^{n} - C_s^{(max)}\right)\right]^2\\ &+\sum_{k=1}^{K}\left[\max\left(0, R_{c,k}^{(min)} - \sum_{s=1}^{S}\phi_{c,s,k}\sum_{n=1}^{N}\psi_{c,s,k}^{n}R_{c,s,k}^{n}\right)\right]^2 + \sum_{k=1}^{K}\left[\max\left(0, \sum_{s=1}^{S}\phi_{c,s,k} - 1\right)\right]^2\\ &+\sum_{s=1}^{S}\sum_{n=1}^{N}\left[\max\left(0, \sum_{k=1}^{K}\psi_{c,s,k}^{n} - 1\right)\right]^2 + \sum_{s=1}^{S}\sum_{k=1}^{K}\sum_{n=1}^{N}\left[\max\left(0, -p_{c,s,k}^{n}\right)\right]^2\\ &+\sum_{s=1}^{S}\sum_{k=1}^{K}\left[\left(\phi_{c,s,k}\right)^2 - \phi_{c,s,k}\right]^2 + \sum_{s=1}^{S}\sum_{k=1}^{K}\sum_{n=1}^{N}\left[\left(\psi_{c,s,k}^{n}\right)^2 - \psi_{c,s,k}^{n}\right]^2\end{aligned} \tag{15}$$

$$\mathbf{X}_i(t+1) = \begin{cases}\mathbf{X}_i^{(SP)}(t) + \beta(t)\left|\mathbf{X}^{(Mean)}(t) - \mathbf{X}_i(t)\right|\cdot\ln\left(\dfrac{1}{\mu_i(t)}\right), & \varepsilon_i(t) > \delta,\\ \mathbf{X}_i^{(SP)}(t) - \beta(t)\left|\mathbf{X}^{(Mean)}(t) - \mathbf{X}_i(t)\right|\cdot\ln\left(\dfrac{1}{\mu_i(t)}\right), & \varepsilon_i(t) \le \delta.\end{cases} \tag{17}$$

solution is far away from the current average position, this candidate solution at the next iteration should accelerate to converge, and vice versa. The term $\mu_i(t)$ is a random variable that prevents the next candidate from falling into a local optimum. In addition to $\mu_i(t)$, $\varepsilon_i(t)$ is also set to reduce the probability that the next candidate solution becomes trapped in local extremes, which can be regarded as the direction for the starting point of the next candidate to search for potential solutions.

3) CONVERGENCE

Ultimately, the proposed TURA algorithm completes once the termination condition is achieved. There are two termination conditions where the first condition is the maximum number of iterations, that is, $t = T$, and the second condition is to reach the stop criterion. The gap ratio $\tau_i(t)$ is defined as

$$\tau_i(t) = \frac{\left| F(\mathbf{X}_i^{(HB)}(t)) - F(\mathbf{X}^{(GB)}(t)) \right|}{F(\mathbf{X}^{(GB)}(t))}, \tag{20}$$

and convergence is achieved if all the gap ratio values $\tau_i(t)$ are smaller than the convergence threshold $F^{(th)}$ as

$$\sum_{i=1}^{I} \mathbb{1}\left(\tau_i(t) \leq F^{(th)}\right) = I. \tag{21}$$

The gap ratio $\tau_i(t)$ in (20) represents the distance between the ith elite candidate solution and the global best solution in iteration t. Equation (21) denotes that the fitness values of all elite candidate solutions are sufficiently close to the global best solution, that is, the fitness function in (14) converges.

The overall procedure for the proposed TURA scheme within a localized C-RAN is presented in Algorithm 1. Moreover, the computational complexity of the proposed TURA algorithm is $\mathcal{O}(I \times T \times S \times K \times N)$, which is linear to the number of candidate solutions, iterations, small cells, users, and subchannels [42]. The performance gain provided by TURA is illustrated via simulation in Section V.

Algorithm 1: Proposed TURA Algorithm

Initialization:
  1) Set the number of candidate solutions I
  2) Set the maximum number of iterations T
  3) Set the iteration counter to $t = 1$
  4) Set the threshold of convergence $F^{(th)}$
  5) Initialize the candidate solution $\mathbf{X}_i(1)$ for all i
  6) Initialize $\mathbf{X}_i^{(HB)}(1) = \mathbf{X}_i(1)$
  7) Initialize $\mathbf{X}^{(GB)}(1)$ by selecting the best solution from $\mathbf{X}_i^{(HB)}(1)$
  8) Initialize CONVERGENCE = **FALSE**
repeat
  Calculate the mean position of candidate solution $\mathbf{X}^{(Mean)}(t)$ by (16)
  Calculate the convergence speed $\beta(t)$ by (19)
  for $i = 1, 2, \ldots, I$ **do**
    Calculate the local attractor $\mathbf{X}_i^{(SP)}(t)$ by (18)
    Update the candidate solution $\mathbf{X}_i(t+1)$ by (16)–(18)
    Calculate the value of fitness function $F(\mathbf{X}_i(t+1))$ by (14)
    if $F(\mathbf{X}_i(t+1)) > F(\mathbf{X}_i(t))$ **then**
      $\mathbf{X}_i^{(HB)}(t+1) = \mathbf{X}_i(t+1)$
    else
      $\mathbf{X}_i^{(HB)}(t+1) = \mathbf{X}_i^{(HB)}(t)$
    end if
    Calculate the value of fitness functions $F(\mathbf{X}_i^{(HB)}(t+1))$ and $F(\mathbf{X}^{(GB)}(t))$ by (14)
    if $F(\mathbf{X}_i^{(HB)}(t+1)) > F(\mathbf{X}^{(GB)}(t))$ **then**
      $\mathbf{X}^{(GB)}(t+1) = \mathbf{X}_i^{(HB)}(t+1)$
    else
      $\mathbf{X}^{(GB)}(t+1) = \mathbf{X}^{(GB)}(t)$
    end if
    Calculate the value $\tau_i(t+1)$ by (20)
  end for
  if $\sum_{i=1}^{I} \mathbb{1}\left(\tau_i(t+1) \leq F^{(th)}\right) = I$ **then**
    CONVERGENCE = **TRUE**
  end if
  Increment of $t \leftarrow t + 1$
until $t = T$ or CONVERGENCE = **TRUE**

IV. PROPOSED CRC GAME AMONG LOCALIZED C-RANs

Considering the hardware limitations, the proposed TURA scheme cannot afford to control a large number of RSCs in a dense SCN. Consequently, the entire network can be viewed as a gathering network of many localized C-RANs, where a CSC will be responsible for the RRM for RSCs within a group. The CRC scheme of subchannel assignment and transmit power allocation for RSCs among localized C-RANs is proposed and described in this section. After eliminating the intra-group interference by resource allocation within each localized C-RAN based on TURA, each CSC conducts a CRC scheme to further alleviate the inter-group interference to achieve higher system performance. Because there is limited and asynchronous information exchanged between the localized C-RANs, it is difficult for all the CSCs to coordinate their RRM under centrally controlled operation. Hence, distributed management is adopted to overcome the above-mentioned difficulty based on game theory, where the resource competition between localized small cells can be formulated as a cooperative game. Each CSC only requires limited information about the probabilities of the chosen actions from other CSCs. We designed a learning algorithm based on the EE of each localized C-RAN for the CSCs to determine their transmit actions, which consist of subchannel assignment and transmit power allocation of their serving RSCs.

A. COOPERATIVE GAME-BASED RESOURCE COMPETITION GAME

In this subsection, the formulation of the proposed CRC scheme among localized small-cell groups is analyzed and

investigated. Based on cooperative game theory, the proposed CRC game in normal form can be denoted as

$$\mathcal{G} = (\mathcal{C}, \{\mathcal{A}_c\}_{c\in\mathcal{C}}, \{U_c\}_{c\in\mathcal{C}}), \tag{22}$$

where the CSC set $\mathcal{C} = \{1, \ldots, C\}$ denotes the set of players in a game, who compete for the subchannel set $\mathcal{N} = \{1, \ldots, N\}$ and the resource for transmitting power of RSCs, that is, the CSCs are the chiefs for action decisions. In addition, $\mathcal{A}_c$ is the action space of CSC c for the power allocation vectors. Let $\kappa^n_{c,s,k} = \psi^n_{c,s,k} \times p^n_{c,s,k}$ be the product indicator of subchannel assignment and allocated transmit power, which determines whether or how much power UE k will be allocated to subchannel n from RSC s in the localized C-RAN c. A finite and discrete action space is a significant requirement for cooperative games. As a result, to form a finite and discrete action space $\mathcal{A}_c$ for each player, let $L_c \in \mathbb{N}$ be the number of discrete powers, and $\kappa^{n,\ell}_{c,s,k}$ represents the ℓth transmit power level from RSC s to UE k over subchannel n, where $n \in \mathcal{N}$, $\ell \in \mathcal{L}_c$ and $\mathcal{L}_c = \{1, \ldots, L_C\}$. Note that if $\psi^n_{c,s,k} = 0$, zero transmit power will be allocated, and $\kappa^{n,\ell}_{c,s,k} = 0$. Thus, we define $\kappa^{0,0}_{c,s,k}$ as the case in which no subchannel is assigned to the user; consequently, the action space of RSC s is expressed as

$$\mathcal{A}_c = \kappa^{0,0}_{c,s,k} \cup \left\{\kappa^{n,\ell}_{c,s,k} : n \in \mathcal{N}, \ell \in \mathcal{L}_C\right\}. \tag{23}$$

We use $\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_C$ to denote the entire action space of all players. Moreover, the last term U_c of (22) is the utility function of CSC c. In game theory, players choose actions based on their own utility functions to obtain the maximum reward. Because the main purpose of this study is to maximize the EE of each localized C-RAN and the performance of the overall system can be further enhanced with the cooperation between CSCs, where each CSC will determine its strategy based on higher EE gain. Accordingly, the utility function for each CSC to decide its transmission strategy for resource allocation can be described as

$$U_c = \tilde{\eta}_c(\mathbf{a}_c), \tag{24}$$

where $\mathbf{a}_c = \{\kappa^{n,\ell}_{c,s,k} | 1 \le s \le S, 1 \le k \le K, 1 \le c \le C, 1 \le n \le N, 0 \le \ell \le L_c, s \in \mathbb{Z}^+, k \in \mathbb{Z}^+, c \in \mathbb{Z}^+, n \in \mathbb{Z}^+, \ell \in \mathbb{Z}^+\}$ represents the vector of actions taken by the CSC c.

B. EXISTENCE OF CE IN PROPOSED SCHEME

In the proposed CRC game $\mathcal{G}$, each CSC aims to maximize its utility function cooperatively and further improve the EE performance of the entire system by choosing an optimal transmission action $\mathbf{a}_c$, which is the solution of (9). Most of the existing works investigate the potential solutions of player strategies by the stable point, which can lead to no play obtaining higher utility gain by changing their determined actions on the Nash equilibrium (NE) [45]. However, a higher utility gain can be acquired by the players cooperatively deciding their strategies via information sharing, which is called a cooperative game. In cooperative games, the stable point that no player will unilaterally deviate from the selected action to others can be held by the CE in game theory. The CE adopted in the proposed CRC scheme is defined as follows.

Definition 1 (Correlated Equilibrium): We denote by $\Delta\mathcal{A}$ the set of all probability distributions over the finite action space $\mathcal{A}$. The probability of the correlated strategy A is given by $P(A)$, where $(P(A))_{A\in\mathcal{A}} \in \Delta\mathcal{A}$. Under the condition that $\forall \mathbf{a}_c \in \mathcal{A}_c$ and $\forall c \in \mathcal{C}$, CSC c will choose strategy $\mathbf{a}_c$ rather than any other strategy $\tilde{\mathbf{a}}_c$ to achieve the stable point of the CE if and only if

$$\sum_{\mathbf{A}_{-c}\in\mathcal{A}_{-c}} P(\mathbf{a}_c, \mathbf{A}_{-c}) \cdot \left[U_c(\mathbf{a}_c, \mathbf{A}_{-c}) - U_c(\tilde{\mathbf{a}}_c, \mathbf{A}_{-c})\right] \ge 0, \tag{25}$$

where $\mathbf{A}_{-c}$ represents the matrix of transmission strategies taken by all CSCs, except for CSC c. In addition, the action space formed by all the CSCs except for CSC c is expressed as $\mathcal{A}_{-c} = \prod_{i\neq c} \mathcal{A}_i$.

It can be observed from (25) that players coordinate their actions with each other by exchanging the probability distribution of strategies cooperatively. In contrast, another well-known concept for analyzing the chosen strategies is NE, where each player is inclined to selfishly decide its actions. Note that NE is a point inside the CE, considering the extreme case in which different strategies are independent. Accordingly, a better overall welfare of the players can be intuitively reached on CE compared with the strategies on NE. This indicates that a higher network EE can be achieved by the CE approach with cooperation between CSCs. The following theorem proves the existence of a CE in our considered SCNs.

Theorem 1: A CE exists in the proposed CRC resource competition game between SCNs.

Proof: Because there is a finite number C of CSCs to choose discrete and finite action spaces in the proposed CRC game $\mathcal{G}$, it is apparent that $\mathcal{G}$ is a finite game. It has been proved by Theorem 1 in [45] that there exists a CE in every finite game. Therefore, the existence of the CE in the proposed CRC game $\mathcal{G}$ can be verified. □

When the subchannels and RSC power are allocated appropriately to the UE, the stable condition, that is, the Pareto optimum, can be reached by the system. Under the stable point of the Pareto optimum, there is no player capable of acquiring a higher reward because it potentially causes losses to others.

Theorem 2: With proper strategies of subchannel assignment and transmit power allocation chosen by the CSCs, the Pareto optimum exists in the proposed CRC resource competition game $\mathcal{G}$.

Proof: Theorem 1 shows that there must exist a CE for the proposed CRC resource competition game $\mathcal{G}$. Moreover, (25) indicates that each player can reach its expected maximum utility in the CE. Hence, the sum of the total expected utilities of all players can be determined by choosing the correlated strategies. The existence of a Pareto optimum can be proved by contradiction as follows: If there is no Pareto optimum in the proposed game $\mathcal{G}$, there must exist

no correlated strategies satisfying (25) and no CE will exist, which obviously contradicts Theorem 1. Consequently, there must exist a Pareto optimum in the proposed game $\mathcal{G}$. □

C. OPERATIONAL PROCESS FOR THE CRC SCHEME

In this subsection, a distributed learning algorithm based on the regret matching procedure [46] is adopted for each CSC to determine its correlated strategy to obtain the CE in the proposed $\mathcal{G}$. Each CSC selects a new strategy of subchannel and power allocation considering the regret value of others not employing strategies in history. In other words, the higher regret value of non-employed strategies indicates a higher probability for the CSC to deviate from its current strategy.

1) INITIALIZATION

The allocation of subchannels, transmit power, and user association will be initially configured from the TURA scheme within the localized small groups. Each CSC calculates its own utility function using (24) from the subchannels, transmit power of the RSC, and user association, which are Φ_c, Ψ_c, and $\mathbf{P}_c$ determined in the operational process of the TURA scheme.

2) EVALUATION OF THE REGRET VALUE

Each CSC determines its potential strategies based on the regret value. Given a history of adopted strategies, the average regret value of CSC c at time t can be expressed as

$$R_c^t(\mathbf{a}_c, \hat{\mathbf{a}}_c) = \max\{D_c^t(\mathbf{a}_c, \hat{\mathbf{a}}_c), 0\}. \tag{26}$$

The R_c^t given in (26) shows that the average regret value depends on the strategy used in the past $\mathbf{a}_c$ and the unused strategy $\hat{\mathbf{a}}_c$, where both $\mathbf{a}_c$ and $\hat{\mathbf{a}}_c \in \mathcal{A}_c$. The term D_c^t represents the quantity of difference between the EE for CSC c at time t to adopt the transmission alignments of $\mathbf{a}_c$ and $\hat{\mathbf{a}}_c$ for subchannel assignment and RSC transmit power allocation. Thus, D_c^t can be formulated as

$$D_c^t(\mathbf{a}_c, \hat{\mathbf{a}}_c) = \frac{1}{t}\sum_{\tau \le t}\left[U_c(\hat{\mathbf{a}}_c, \mathbf{A}_{-c}^{\tau}) - U_c^{\tau}(\mathbf{A}^{\tau})\right], \tag{27}$$

where $\mathbf{A}^{\tau}$ is the total action taken by all CSCs. Note that superscript of τ indicates the instant time less than time t.

3) PROBABILITY EXCHANGING AND STRATEGY UPDATING

CSCs will choose strategies with the highest profit of utility. After evaluating R_c^t of potential strategies, the probabilities that the strategies are prone to be selected are updated as follows:

$$\mathbf{w}_c^t(\hat{\mathbf{a}}_c^t) = \begin{cases} \frac{1}{\xi} R_c^t(\mathbf{a}_c^{\tau}, \hat{\mathbf{a}}_c^t), & \hat{\mathbf{a}}_c^t \neq \mathbf{a}_c^{\tau}, \\ 1 - \sum\limits_{\hat{\mathbf{a}}_c^t \in \mathcal{A}_c} \frac{1}{\xi} R_c^t(\mathbf{a}_c^{\tau}, \hat{\mathbf{a}}_c^t), & \hat{\mathbf{a}}_c^t = \mathbf{a}_c^{\tau}, \end{cases} \tag{28}$$

where $\mathbf{a}_c^{\tau}$ is the previously strategy used at τ, whilst the unused strategy at t is denoted as $\hat{\mathbf{a}}_c^t$. The term ξ in (28) is a normalization constant that is a non-negative and sufficiently large number to normalize the summation of probabilities for different strategies to one. Each CSC exchanges its probabilities of potential strategies with each other. However, errors may occur due to asynchronous information exchange from other CSCs. Therefore, the exchange probability vector, as shown in Fig. 3 can be represented as

$$\tilde{\mathbf{w}}_{-c}^t(\hat{\mathbf{a}}_{-c}^t) = \mathbf{w}_{-c}^t(\hat{\mathbf{a}}_{-c}^t) \cdot (1 + \rho \Delta h), \tag{29}$$

where $\mathbf{w}_{-c}^t$ is the set for the probability of potential strategies from all the CSCs, except for CSC c. Moreover, the parameter $\rho \in [0, 1]$ is the error ratio due to asynchronous exchanged information and $\Delta h \sim \mathcal{CN}(0, 1)$. As a result, each CSC can determine its strategy for transmission alignment as

$$\mathbf{a}_c^{t+1} = \arg\max_{\hat{\mathbf{a}}_c^t \in \mathcal{A}_c} \mathbf{w}_c^t(\hat{\mathbf{a}}_c^t). \tag{30}$$

4) CONVERGENCE

The ultimately determined strategy for each CSC is achieved according to the following criterion:

$$\frac{|U_c^{t+1} - U_c^t|}{U_c^t} \le \theta^{(th)}, \qquad \forall c. \tag{31}$$

The subchannel assignment and transmit power of RSCs will be allocated by their serving CSC within the localized C-RAN when the criteria of (31) meet. Otherwise, the CSCs continue to perform the algorithm until (31) is satisfied. The detailed procedure for the proposed CRC algorithm is described in Algorithm 2.

Algorithm 2: Proposed CRC Algorithm

1: **Input**: Channel condition $g_{c,s,k}^n, \forall c, \forall s, \forall k, \forall n$, parameters ξ, threshold $\theta^{(th)}$
2: **for** $c = 1, \ldots, C$ **do**
3: Set the initial iteration as $t = 1$
4: CSC c initializes its strategy $\mathbf{a}_c^1$ of subchannel assignment and transmit power allocation for the users according to Φ_c, Ψ_c, $\mathbf{P}_c$ in TURA scheme
5: **repeat**
6: Evaluate the average regret value of different strategies based on its own utility function of EE by calculating (26)
7: Update the probability for the strategy according to (28)
8: Exchange the probability of potential strategies with other CSCs according to (29)
9: Choose the strategy for iteration $t + 1$ given the probability distribution of $\mathbf{w}_c^t(\hat{\mathbf{a}}_c^t)$ by (30)
10: $t = t + 1$
11: **until** Convergence: $\frac{|U_c^{t+1} - U_c^t|}{U_c^t} \le \theta^{(th)}$
12: **end for**

V. PERFORMANCE EVALUATIONS

In this section, the performance evaluation of the proposed HARM, TURA, and CRC is provided through simulations.

TABLE 2. Main system parameters.

Parameters of the system	Value
Number of subchannels N	50
Subchannel bandwidth W	360 kHz
Carrier frequency	2.6 GHz
Noise power N_0	-174 dBm/Hz
Circuit power $P^{(\mathrm{CA})}/P^{(\mathrm{CS})}$	6.8/4.3 W
Signal power overhead $P^{(\mathrm{O})}$	1 dBm
Maximum transmit power $P^{(\max)}$	20 dBm
Minimum data rate requirement $R^{(\min)}$	10 Mbps
Coverage area	90×90 (m^2)
Parameters of HARM scheme	**Value**
Maximum number of iterations T	600
Number of candidate solutions I	40
Penalty factor α	1.5
Search range $\beta^{(\max)}/\beta^{(\min)}$	1.2/0.5
Update threshold δ	0.5
Convergence threshold $F^{(\mathrm{th})}/\theta^{(\mathrm{th})}$	$10^{-4}/10^{-4}$

Consider an SCN, the RSCs are deployed in a grid in a square coverage, and users are uniformly distributed within the coverage area. Each user was located at a minimum distance of 2 m from each RSC. The network channel model and system parameters are considered based on [47], [48], where the case of an indoor dense urban information society in [48] is taken as the reference. Channel fading is composed of path loss and Rayleigh fading. The path loss model between RSCs and users depends on sight conditions such as line-of-sight (LoS) and non-line-of-sight (NLoS), that is, $PL = 18.7\log_{10}(d_{s,k}) + 46.8 + 174.32$ in LoS and $PL = 36.8\log_{10}(d_{s,k}) + 43.8 + 174.32 + 5(n_w - 1)$ in NLoS, where $d_{s,k}$ is the distance between RSC s and user k in meters, and n_w is a random variable regarding the number of walls between the transmitter and receiver. Rayleigh fading is modeled as an independent and identically distributed Gaussian distribution random variable. Note that the penalty weight of $\alpha = 1.5$ is empirically defined to leverage fitness function with both EE objective and satisfaction of constraints. The main parameter settings are listed in Table 2. Moreover, if not mentioned specifically, the data rate of each user, maximum transmit power of the RSC, and fronthaul capacity of the RSC are set to be identical to $R_{c,k}^{(min)} = R^{(min)}$, $P_{c,s}^{(max)} = P^{(max)}$ and $B_{c,s}^{(max)} = B^{(max)}$, respectively, $\forall c, k, s$. All simulation results were averaged over random user locations and channel conditions according to the Monte Carlo runs.

A. CONVERGENCE OF THE PROPOSED CRC SCHEME

In this subsection, the convergence of the proposed CRC scheme is presented in Fig. 4 under different numbers of users K. Note that the CRC scheme mentioned in Section IV is designed to perform resource competition and interference management among the six C-RANs in the simulations. It can be observed from Fig. 4 that the CRC algorithm can quickly converge to its CE on EE performance under several iterations by adopting the regret-based learning algorithm. More iterations are required to achieve the CE in the proposed CRC scheme with a higher number of serving users K owing to more potential strategies being searched for resource allocation.

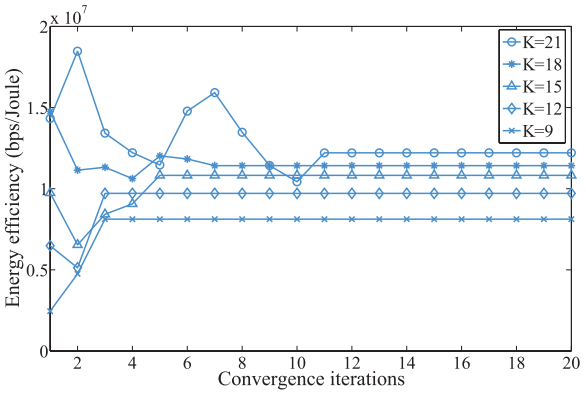


FIGURE 4. Energy efficiency versus number of iterations in proposed CRC scheme under $C = 6$ and $B^{(max)} = 20$ Mbps.

B. PERFORMANCE OF THE PROPOSED TURA SCHEME

In this subsection, the simulation results are provided to demonstrate the effectiveness of the proposed TURA scheme on traffic control, the effects of the RSC on/off mechanism, and the capacity limitation of fronthaul. Furthermore, the effect of signaling overhead for traffic control on EE performance is also discussed. Because the effect of traffic control is a critical concern in this study, a benchmark of RSRP-based user association and resource allocation (RURA) method is considered for performance comparison, where it adopts the identical power and subchannel allocations as TURA along with RSRP-based user association. Here, the probability of outage is calculated as the ratio of the number of UEs with unsatisfied QoS to that of all serving UEs, i.e.,

$$\Delta_{\mathrm{out}} = \frac{1}{CK}\sum_{k=1}^{K}\sum_{c=1}^{C}\mathbb{1}\left(\sum_{s=1}^{S}\phi_{c,s,k}\sum_{n=1}^{N}\psi_{c,s,k}^{n}R_{c,s,k}^{n} < R_{c,k}^{(\min)}\right), \tag{32}$$

where the indicator function is regarded as unsatisfaction of QoS in constraint (9c). Moreover, we define the probability of user offloading as the portion of the number of UEs with the alternated policy from the original RSRP-based association. That is, the connection changes from the strongest RSC to a compromised one with offloading capability achieving higher EE performance.

1) USER OFFLOADING AND TRAFFIC CONTROL OF THE TURA SCHEME

Fig. 5 depicts the relationship between QoS satisfaction and traffic control effect in terms of outage probability and traffic control probability, respectively. The capacities of fronthaul are set as $B^{(max)} = \{10, 10, 20, 10\}$ Mbps for the four RSCs. A user is associated with each RSC at the beginning, and then other users are added into the network at a later time, i.e., $K = \{4, 5, 6, 7\}$. The left subplot of Fig. 5 shows that there is no outage caused in both TURA and RURA schemes with $K = 4$ because the capacity of fronthaul is sufficient for the RSC to satisfy the QoS of each user. Furthermore, the outage probability is lower in the proposed TURA scheme compared with the RURA method with increasing number of users in

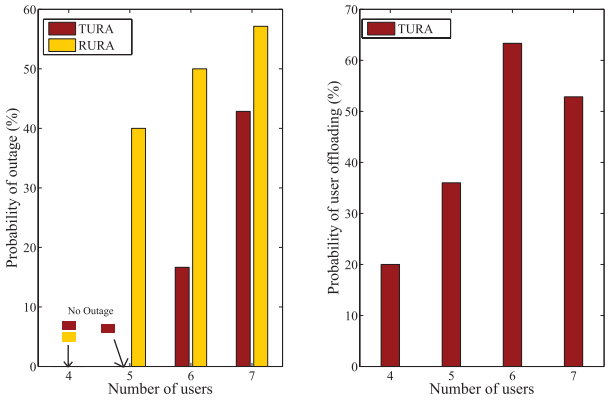


FIGURE 5. Traffic control effect in terms of the probability of outage (left) and user offloading (subplot) versus different numbers of users under $S = 4$ RSCs and $B^{(max)} = \{10, 10, 20, 10\}$ Mbps.

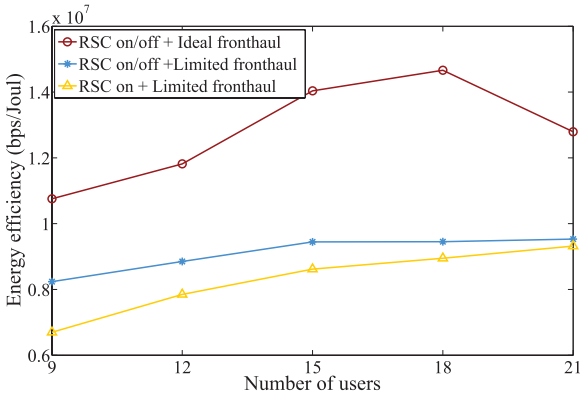


FIGURE 6. Performance comparison for three scenarios in proposed TURA scheme in terms of EE versus number of users under $S = 6$ and $B^{(max)} = 20$ Mbps.

the network, which illustrates the merits of the TURA scheme on user offloading as shown in the right subplot to overcome the restriction of insufficient fronthaul capacity for RSCs. Note that the probability of user offloading control decreases from $N = 6$ to $N = 7$ when the network traffic is overloaded. This is because the total capacity of fronthaul is limited, and there is no appropriate configuration of user association to satisfy the data rate requirements of users. Therefore, Fig. 5 demonstrates the traffic control capability of improving QoS satisfaction with the limitation of fronthaul capacity.

2) EFFECT OF RSC ON/OFF MECHANISM AND FRONTHAUL CAPACITY LIMIT

In Fig. 6, the relationship between EE and different users $K = \{9, 12, 15, 18, 21\}$ under $S = 6$ and $B^{(max)} = 20$ Mbps is illustrated for three scenarios. The scenario of RSC on/off with unlimited fronthaul capacity provides a better performance on EE than the scenario of RSC on/off with limited fronthaul because the data rate of the user is not restricted by the capacity of the fronthaul. In addition, the performance comparison between the scenarios of RSC on/off with limited fronthaul and RSC on with limited fronthaul depicts the merit of the RSC on/off mechanism on the power saving of RSCs

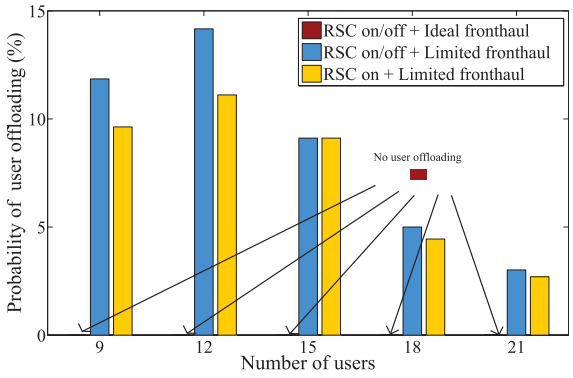


FIGURE 7. Performance comparison for three scenarios in the proposed TURA scheme in terms of the probability of user offloading versus number of users under $S = 6$ and $B^{(max)} = 20$ Mbps.

to enhance network EE. Furthermore, it can be observed that the effectiveness of the mechanism of the RSC on/off mechanism is revealed, especially under lower traffic loads, that is, $K = 9$, where the RSCs tend to be turned off to conserve energy. Fig. 7 illustrates the probability of user offloading under different numbers of users for the proposed TURA scheme. It can be observed that there is no traffic control performed in the scenario of RSC on/off with ideal fronthaul because sufficient capacity is available to support the required data rate for the fronthaul. The RSC on/off mechanism affects traffic control because power can be conserved by offloading the user to other RSCs and turning off the RSC that is no longer serving users. Given that a smaller number of users will lead to a higher tendency for the RSCs to be turned off by conducting traffic control, the effect of traffic control on energy conservation for RSC power is revealed, especially under lower network traffic loads, for example, $K = \{9, 12\}$. It can be concluded that the effect of traffic control is mainly influenced by the limitation of fronthaul capacity.

Fig. 8 illustrates the probability of user offloading and EE performance under different signaling power of $P^{(O)} = \{0, 1, 10, 20, 35\}$ dBm for proposed TURA scheme with the number of RSC $S = 6$ and $B^{(max)} = 20$ Mbps. Note that the signaling power $P^{(O)}$ is considered in the proposed TURA scheme to reduce the ping-pong effect in the handover process as mentioned in (6). It can be seen that both the probability of user offloading and EE decrease with the increased signaling power. The reason is that traffic control is performed either for QoS satisfaction when the fronthaul link is overloaded or for energy conservation. Hence, the network EE decreases with the increasing signaling power owing to user offloading to achieve the minimum data rate requirements.

C. PERFORMANCE OF THE PROPOSED HARM SCHEME

The performance comparison of the EE and outage probability of the proposed HARM scheme considering the effects of the number of localized C-RANs and the error ratio of information exchanged ρ versus the number of users is shown in Figs. 9 and 10, respectively. In Fig. 9, the EE performance of HARM with zero error ratio $\rho = 0$ degrades

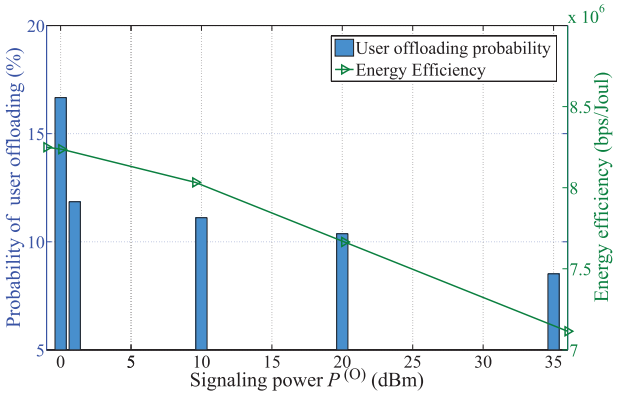


FIGURE 8. EE and probability of user offloading versus signaling power $P^{(O)}$ with the proposed TURA scheme under $S = 6$ and $B^{(max)} = 20$ Mbps.

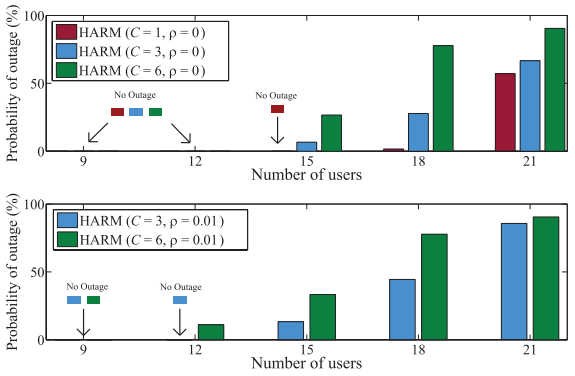


FIGURE 10. Performance comparison of the proposed HARM scheme in terms of the probability of outage versus number of users under $S = 6$, $B^{(max)} = 30$ Mbps, $C = \{1, 3, 6\}$, and $\rho = \{0, 0.01\}$.

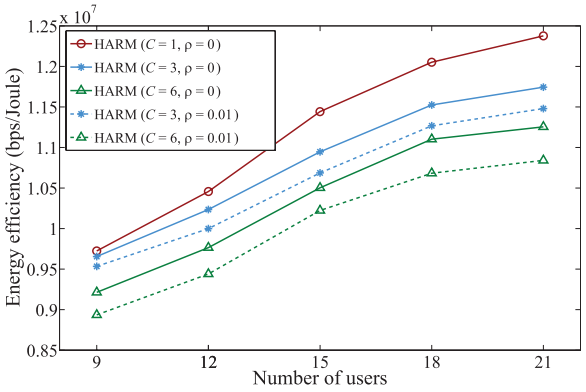


FIGURE 9. Performance of the proposed HARM scheme in terms of EE versus number of users under $S = 6$, $B^{(max)} = 30$ Mbps, $C = \{1, 3, 6\}$, and $\rho = \{0, 0.01\}$.

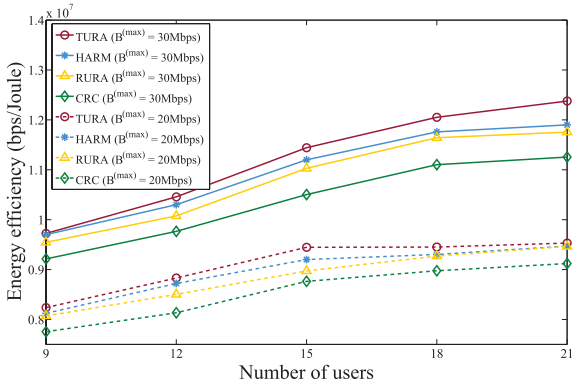


FIGURE 11. Performance comparison between proposed TURA, HARM, RURA, and CRC schemes in terms of energy efficiency versus number of users under $S = 6$ and $B^{(max)} = \{20, 30\}$ Mbps.

with increasing number of localized C-RANs from $C = 1$ to $C = 6$. This is because more localized C-RANs that divide the network will provoke a comparably more distributed control for resource allocation. Consequently, the coordination among localized C-RANs cannot be performed simultaneously under $C = 6$ C-RANs compared with the totally centralized control under $C = 1$. With a higher error ratio of $\rho = 0.01$, the HARM has a lower EE owing to asynchronous exchanged information and greater power consumption under a large-scale network. Furthermore, as shown in Fig. 10, the outage probability of the proposed HARM scheme with zero error ratio $\rho = 0$ asymptotically increases with the increment of localized C-RANs owing to non-coordinated interference management and restricted traffic control. Under a higher error ratio of asynchronous information of $\rho = 0.01$, it potentially results in a higher probability of erroneous resource allocation strategies exchanged between CSCs, which induces inappropriate interference management. In other words, RSCs in the associated localized C-RAN cannot alleviate interference by allocating proper subchannels and transmit power, which also reflects the degradation of EE performance from $\rho = 0$ to $\rho = 0.01$, as illustrated in Fig. 9.

D. PERFORMANCE COMPARISONS WITH EXISTING BENCHMARKS

In this subsection, we evaluate the performance of the proposed HARM scheme to observe the integrated effects of both the TURA and CRC methods. Note that the centralized control of the proposed TURA scheme considers that all RSCs are managed by a single CSC with upper MAC functions, whereas the lower MAC functions reside in RSCs with confined fronthaul capacity. Moreover, the performance of HARM was investigated along with the proposed CRC scheme employed in distributed small cells. Note that the RURA benchmark adopts an RSRP-based user association along with non-adjustable power and subchannel allocations.

Fig. 11 illustrates EE performance of the proposed TURA, HARM, and CRC schemes compared with the benchmark of RURA under different users $K = \{9, 12, 15, 18, 21\}$ and fronthaul capacities $B^{(max)} = \{20, 30\}$ Mbps. It can be observed that higher EE is reached with sufficient fronthaul capacity of $B^{(max)} = 30$ Mbps. The performance of EE in TURA scheme outperforms the other methods in both scenarios owing to the centralized control by CSC, which strikes a compelling balance of interference management and traffic control under limited fronthaul links. The EE performance of HARM is slightly degraded from TURA

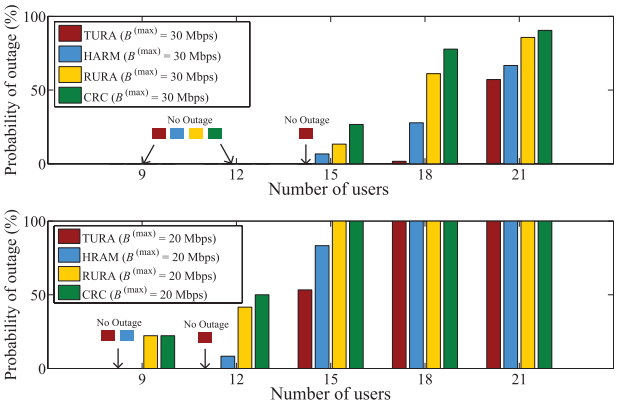


FIGURE 12. Performance comparison between the proposed TURA, HARM, RURA, and CRC schemes in terms of outage probability versus number of users under $S = 6$ and $B^{(max)} = \{20, 30\}$ Mbps.

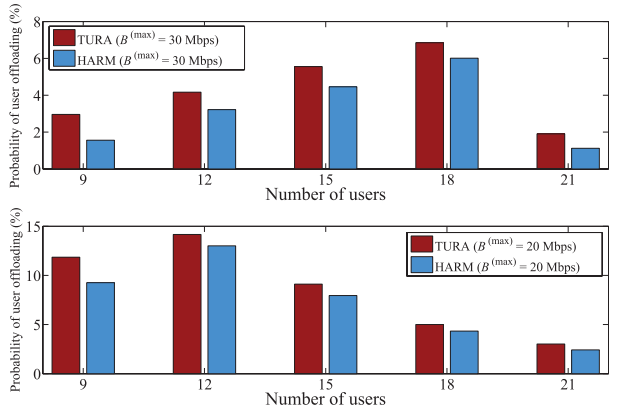


FIGURE 13. Probability of user offloading in the proposed TURA and HARM scheme versus number of users under $S = 6$ and $B^{(max)} = \{20, 30\}$ Mbps.

because the interference among localized RSCs is not as well alleviated as that performed in TURA. Moreover, the users cannot be readily offloaded between different RSCs owing to the restricted capacity of fronthaul. In addition, higher EE is acquired by RURA than by CRC which is induced by more existing interferences from asynchronous information exchanged and inappropriate resource configuration.

The outage probability considering different fronthaul capabilities is illustrated in Fig. 12, whereas Fig. 13 depicts the probability of user offloading with the TURA and HARM schemes under different numbers of users $K = \{9, 12, 15, 18, 21\}$ and fronthaul capacities $B^{(max)} = \{20, 30\}$ Mbps. Note that there is no traffic control mechanism in the compared schemes of RURA and CRC. It can be inferred that the proposed TURA scheme outperforms the benchmarks of RURA and CRC owing to the centralized configuration of resource management and traffic control. On the other hand, the proposed HARM has a slightly higher outage probability and a lower tendency of traffic control because the users are not frequently offloaded between two neighboring RSCs. We can also infer that the outage probability can be substantially reduced with a higher fronthaul capacity, that is, a lower outage is achieved under $B^{(max)} = 30$ Mbps compared with that under $B^{(max)} = 20$ Mbps. However, users suffer from full outage among all schemes because of the higher number of users of $K = \{18, 21\}$ and insufficient fronthaul of $B^{(max)} = 20$ Mbps. Moreover, Fig. 13 depicts that the increased offloading probability is revealed owing to asymptotically saturated fronthaul. The probability of user offloading starts to decrease from $K = 18$ under $B^{(max)} = 30$ and from $K = 12$ under $B^{(max)} = 20$ because the fronthaul is overloaded, which is incapable of supporting the data rate requirements. In addition to QoS satisfaction, user offloading will also be performed to conserve energy, resulting in potential sleep-mode RSCs.

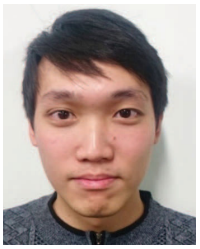
VI. CONCLUSION

In this paper, we have proposed hybrid controlled user association and resource management to resolve the problem of subchannel and transmit power allocation for EE maximization considering limited fronthaul capacity under large-scale green C-RANs. Within a localized C-RAN, the CSC centrally performs the proposed TURA scheme for the RSCs to mitigate the intra-group interference and tackle the issue of limited fronthaul capacity for user QoS requirements. Furthermore, the proposed CRC scheme can analytically achieve CE and alleviate inter-group interference among localized C-RANs. The Pareto optimum of the CRC scheme has been theoretically proven based on the game theory. Moreover, the simulation results have revealed the effectiveness of traffic control by verifying the convergence of the CRC scheme. In addition, the EE performance of the proposed TURA scheme outperforms the other existing methods owing to fully centralized management under the consideration of an ideal fronthaul. Despite the slightly lower EE performance of the proposed HARM scheme compared with the ideal case of TURA, HARM is capable of sustaining an appropriate EE compared with existing schemes under feasible implementation of practical green RANs.

REFERENCES

[1] F. Boccardi, R. W. Heath, Jr., A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
[2] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3GPP heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 10–21, Jun. 2011.
[3] T. Nakamura, S. Nagata, A. Benjebbour, Y. Kishiyama, T. Hai, S. Xiaodong, Y. Ning, and L. Nan, "Trends in small cell enhancements in LTE advanced," *IEEE Commun. Mag.*, vol. 51, no. 2, pp. 98–105, Feb. 2013.
[4] E. Oh, B. Krishnamachari, X. Liu, and Z. Niu, "Toward dynamic energy-efficient operation of cellular network infrastructure," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 56–61, Jun. 2011.
[5] J. B. Rao and A. O. Fapojuwo, "A survey of energy efficient resource management techniques for multicell cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 154–180, 1st Quart., 2014.
[6] R. Mahapatra, Y. Nijsure, G. Kaddoum, N. U. Hassan, and C. Yuen, "Energy efficiency tradeoff mechanism towards wireless green communication: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 686–705, 1st Quart., 2016.
[7] L.-H. Shen, S.-F. Hung, and K.-T. Feng, "Hybrid energy resource allocation for simultaneous wireless information and power transfer for green communications," *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 4, pp. 2121–2138, Dec. 2021.

[8] A. G. White, "GreenTouch final results from green meter research study—Reducing the net energy consumption in communications networks by up to 98% by 2020," GreenTouch White Paper, Version 2.0, Aug. 2015.
[9] Q. Wang, Z. Gao, Y. Xu, and H. Xie, "Energy-efficient optimization for IRS-assisted wireless-powered communication networks," in *Proc. IEEE 93rd Veh. Technol. Conf. (VTC-Spring)*, Apr. 2021, pp. 1–5.
[10] D. Jiang, Z. Wang, W. Wang, Z. Lv, and K.-K.-R. Choo, "AI-assisted energy-efficient and intelligent routing for reconfigurable wireless networks," *IEEE Trans. Netw. Sci. Eng.*, early access, Apr. 27, 2021, doi: 10.1109/TNSE.2021.3075428.
[11] S. H. Lee, M. Kim, H. Shin, and I. Lee, "Belief propagation for energy efficiency maximization in wireless heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 56–68, Jan. 2021.
[12] W. Liu, S. Han, C. Yang, and C. Sun, "Massive MIMO or small cell network: Who is more energy efficient?" in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, Apr. 2013, pp. 24–29.
[13] C.-H. Fang, P.-R. Li, and K.-T. Feng, "Joint interference cancellation and resource allocation for full-duplex cloud radio access networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 6, pp. 3019–3033, Jun. 2019.
[14] Z. Q. Al-Abbasi, K. M. Rabie, and D. K. C. So, "EE optimization for downlink NOMA-based multi-tier CRANs," *IEEE Trans. Veh. Technol.*, vol. 70, no. 6, pp. 5880–5891, Jun. 2021.
[15] L. Ferdouse, I. Woungang, A. Anpalagan, and S. Erkucuk, "Energy efficient downlink resource allocation in cellular IoT supported H-CRANs," *IEEE Trans. Veh. Technol.*, vol. 70, no. 6, pp. 5803–5816, Jun. 2021.
[16] S. O. Oladejo, S. O. Ekwe, and L. A. Akinyemi, "Multi-tier multi-domain network slicing: A resource allocation perspective," in *Proc. IEEE Africon Conf. (AFRICON)*, Sep. 2021, pp. 1–6.
[17] S. O. Oladejo and O. E. Falowo, "Latency-aware dynamic resource allocation scheme for multi-tier 5G network: A network slicing-multitenancy scenario," *IEEE Access*, vol. 8, pp. 74834–74852, 2020.
[18] L. M. P. Larsen, A. Checko, and H. L. Christiansen, "A survey of the functional splits proposed for 5G mobile crosshaul networks," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 146–172, 1st Quart, 2019.
[19] J. Hoydis, M. Kobayashi, and M. Debbah, "Green small-cell networks," *IEEE Veh. Technol. Mag.*, vol. 6, no. 1, pp. 37–43, Mar. 2011.
[20] *C-RAN: The Road Towards Green RAN*, China Mobile, Beijing, China, 2011.
[21] Y.-T. Huang, C.-H. Fang, L.-H. Shen, and K.-T. Feng, "Optimal functional split for processing sharing based CoMP for mixed eMBB and uRLLC traffic," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2020, pp. 1–6.
[22] C.-H. Fang, L.-H. Shen, T.-P. Huang, and K.-T. Feng, "Delay-aware admission control and beam allocation for 5G functional split enhanced millimeter wave wireless fronthaul networks," *IEEE Trans. Wireless Commun.*, early access, Sep. 21, 2021, doi: 10.1109/TWC.2021.3112299.
[23] J. Ghimire and C. Rosenberg, "Impact of limited backhaul capacity on user scheduling in heterogeneous networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2014, pp. 2480–2485.
[24] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in multi-cell OFDMA systems with limited backhaul capacity," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3618–3631, Oct. 2012.
[25] K. Zeb, X. Zhang, and Z. Lu, "High capacity mode division multiplexing based MIMO enabled all-optical analog millimeter-wave over fiber fronthaul architecture for 5G and beyond," *IEEE Access*, vol. 7, pp. 89522–89533, 2019.
[26] A. Younis, T. X. Tran, and D. Pompili, "Energy-efficient resource allocation in C-RANs with capacity-limited fronthaul," *IEEE Trans. Mobile Comput.*, vol. 20, no. 2, pp. 473–487, Feb. 2021.
[27] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. G. Andrews, "User association for load balancing in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2706–2716, Jun. 2013.
[28] Q. Chen, G. Yu, R. Yin, and G. Y. Li, "Joint user association and resource allocation for energy-efficient multi-stream aggregation," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 2482–2487.
[29] A. Khanafer, W. Saad, and T. Basar, "Context-aware wireless small cell networks: How to exploit user information for resource allocation," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 3341–3346.
[30] A. A. Zahrani and F. R. Yu, "An energy-efficient resource allocation and interference management scheme in green heterogeneous networks using game theory," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5384–5396, Jul. 2016.
[31] C. G. Diaz, B. Maham, D. Niyato, and A. S. Madhukumar, "Strategic device-to-device communications in backhaul-constrained wireless small cell networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2014, pp. 1661–1666.
[32] M. Bennis, S. M. Perlaza, and M. Debbah, "Learning coarse correlated equilibria in two-tier wireless networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2012, pp. 1592–1596.
[33] D. Wu, L. Zhou, and Y. Cai, "Energy-efficient resource allocation for uplink orthogonal frequency division multiple access systems using correlated equilibrium," *IET Commun.*, vol. 6, no. 6, pp. 659–667, Jun. 2012.
[34] P. Semasinghe, E. Hossain, and K. Zhu, "An evolutionary game for distributed resource allocation in self-organizing small cells," *IEEE Trans. Mobile Comput.*, vol. 14, no. 2, pp. 274–287, Feb. 2015.
[35] H. Zhang, C. Jiang, N. C. Beaulieu, X. Chu, X. Wen, and M. Tao, "Resource allocation in spectrum-sharing OFDMA femtocells with heterogeneous services," *IEEE Trans. Commun.*, vol. 62, no. 7, pp. 2366–2377, Jul. 2014.
[36] N. Prasad, M. Arslan, and S. Rangarajan, "Exploiting cell dormancy and load balancing in LTE HetNets: Optimizing the proportional fairness utility," *IEEE Trans. Commun.*, vol. 62, no. 10, pp. 3706–3722, Oct. 2014.
[37] A. Abdelnasser and E. Hossain, "On resource allocation for downlink power minimization in OFDMA small cells in a cloud-RAN," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2014, pp. 1–6.
[38] Y. Jin, L. Qiu, and X. Liang, "Small cells on/off control and load balancing for green dense heterogeneous networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2015, pp. 1530–1535.
[39] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[40] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in OFDMA systems with large numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3292–3304, Sep. 2012.
[41] F. van den Bergh and A. P. Engelbrecht, "A new locally convergent particle swarm optimiser," in *Proc. IEEE Int. Conf. Syst., Man Cybern. (SMC)*, Oct. 2002, pp. 96–101.
[42] J. Sun, W. Fang, X. Wu, V. Palade, and W. Xu, "Quantum-behaved particle swarm optimization: Analysis of individual particle behavior and parameter selection," *Evol. Comput.*, vol. 20, no. 3, pp. 349–393, Sep. 2012.
[43] Y. Zhao, X. Li, Y. Li, and H. Ji, "Resource allocation for high-speed railway downlink MIMO-OFDM system using quantum-behaved particle swarm optimization," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2013, pp. 2343–2347.
[44] Q. Xu, X. Li, and H. Ji, "Multiple resource allocation in OFDMA downlink networks: End-to-end energy-efficient approach," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2014, pp. 3957–3962.
[45] S. Hart and D. Schmeidler, "Existence of correlated equilibria," *Math. Oper. Res.*, vol. 14, no. 1, pp. 18–25, Feb. 1989.
[46] S. Hart and A. Mas-Colell, "A simple adaptive procedure leading to correlated equilibrium," *Econometrica*, vol. 68, no. 5, pp. 1127–1150, Sep. 2000.
[47] *Evolved Universal Terrestrial Radio Access (E-UTRA); Further Advancements for E-UTRA Physical Layer Aspects*, document 3GPP TR 36.814, Version 9.0.0, Mar. 2010.
[48] *Deliverable D6.1 Simulation Guidelines, METIS 2020 Project ICT- 317669-METIS/D6.1*, METIS, Oct. 2013. [Online]. Available: https://metis2020.com/wp-content/uploads/deliverables/METIS_D6.1_v1.pdf



LI-HSIANG SHEN (Member, IEEE) received the Ph.D. degree from the Institute of Communication Engineering, NCTU, Hsinchu, Taiwan, in 2020. Since January 2021, he has been a Postdoctoral Researcher with the Department of Electrical and Computer Engineering (ECE), National Yang Ming Chiao Tung University (NYCU), Hsinchu. In 2018 and 2019, he was a Visiting Scholar with the Next Generation Wireless (NGW) Research Group, Electrical and Computer Engineering Department, University of Southampton, Southampton (Soton), U.K. His research interests include millimeter-wave/terahertz interference management and resource allocation, mobility-aware beam control protocol for wireless local area networks (WLANs), new radio (NR), the Internet of Things (IoT), unmanned aerial vehicle (UAV) communications, and machine/deep learning empowered wireless networks. He was awarded the first prize of the Broadcom Foundation Asia Pacific Workshop. In 2021, he was also awarded the IEEE Best Ph.D. Thesis Award, NYCU Outstanding Graduate Student—Ph.D. Research, and Phi Tau Phi Scholastic Honor Society of the Republic of China (R.O.C.).

CHIA-LIN TSAI received the M.S. degree from the Department of Institute of Communication Engineering, National Chiao Tung University (NCTU), Hsinchu, in 2016. His research interests include small-cell green communications, interference management, load-balancing mechanisms, and radio resource management.

CHIA-YU WANG received the M.S. degree from the Department of Institute of Communication Engineering, National Chiao Tung University (NCTU), Hsinchu, in 2016. Her research interests include small-cell wireless networks, interference management, load balancing, and radio resource allocation.

KAI-TEN FENG (Senior Member, IEEE) received the B.S. degree from the National Taiwan University, Taipei, Taiwan, in 1992, the M.S. degree from the University of Michigan, Ann Arbor, in 1996, and the Ph.D. degree from the University of California at Berkeley, Berkeley, in 2000. Since August 2011, he has been a Full Professor with the Department of Electrical and Computer Engineering, National Yang Ming Chiao Tung University (NYCU), Hsinchu, Taiwan, where he was an Associate Professor and an Assistant Professor, from August 2007 to July 2011 and from February 2003 to July 2007, respectively. He worked as the Associated Dean of the Electrical and Computer Engineering College, NYCU, starting from February 2017. From July 2009 to March 2010, he was a Visiting Research Fellow with the Department of Electrical and Computer Engineering, University of California at Davis. Between 2000 and 2003, he was an In-Vehicle Development Manager/a Senior Technologist at OnStar Corporation, a subsidiary of General Motors Corporation, where he worked on the design of future telematics platforms and in-vehicle networks. His current research interests include broadband wireless networks, cooperative and cognitive networks, wireless indoor localization and tracking, and device-free wireless sensing technologies. He received the Best Paper Award from the Spring 2006 IEEE Vehicular Technology Conference, which ranked his paper first among the 615 accepted papers. He also received the Outstanding Youth Electrical Engineer Award in 2007 from the Chinese Institute of Electrical Engineering and the Distinguished Researcher Award from NCTU in 2008, 2010, and 2011. He has been serving as a Technical Advisor for IEEE-HKN Honor Society and the National Academy of Engineering (NAE) Grand Challenges Scholars Program (GCSP) at NCTU, since 2018. He has also served technical program committees at various international conferences.